



C I N D E: R e p o r t e v a r i a b l e d e p e n d i e n t e

— Martes 14 de marzo —

Contenido

1. Objetivo del proyecto
2. Categorías
3. Introducción a los gráficos
4. Análisis de los resultados



Objetivo del proyecto

Apoyar en el proceso de identificación de jóvenes en situación de vulnerabilidad mediante la analítica de datos y el desarrollo de un modelo de machine learning

$$y = f(X)$$

y es la variable que
queremos predecir



Categorías de vulnerabilidad

Categorías

Categoría	
Condiciones de vida digna	Normalización de las violencias
Dinámicas familiares	Situaciones de ansiedad, depresión, estrés post traumático y efectos sobre autoestima
Dinámicas territoriales y familia	Situaciones de victimización por intimidación
Intimidaciones por parte de respondientes	Violencia de Género contra la mujer
Percepción de violencias basadas en género contra personas LGBTI+	

*No se tendrá en cuenta la categoría de violencias basadas en género contra personas LGBTI+ porque hasta el momento solo tenemos 3 estudiantes que se reconocen como LGBTI+, lo anterior son muy pocas observaciones para incluir en el modelo



Ejemplo

Situaciones de victimización por intimidación	Siempre	Casi Siempre	Casi Nunca	Nunca
No me dejan participar, me excluyen	6	4	2	0
Me obligan a hacer cosas peligrosas para mí	12	8	4	0
Rompen mis cosas a propósito	6	4	2	0
Me insultan	6	4	2	0
Me pegan coscorriones, puñetazos, patadas	12	8	4	0
Me chiflan, me gritan	6	4	2	0
Dicen a otros que no estén conmigo o que no me hablen	6	4	2	0
Me desprecian	6	4	2	0
Me amenazan para que haga cosas que no quiero	12	8	4	0
Me obligan a hacer cosas que están mal	12	8	4	0



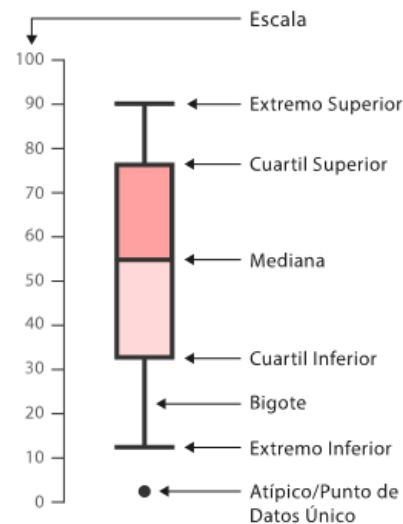
Introducción a gráficos utilizados

Definición: Diagrama de cajas y bigotes

Un gráfico de cajas y bigotes, también conocido como diagrama de caja, es una representación visual de un conjunto de datos numéricos a través de cinco medidas estadísticas: **el valor mínimo, el primer cuartil (Q1), la mediana (Q2), el tercer cuartil (Q3) y el valor máximo.**

El gráfico consiste en una caja que va **desde el primer cuartil hasta el tercer cuartil, con una línea que representa la mediana en el medio.** Los "bigotes" se extienden desde la caja hasta el valor mínimo y máximo, y se pueden incluir puntos individuales fuera de los bigotes para mostrar valores atípicos.

El gráfico de cajas y bigotes es **útil para identificar la dispersión y la simetría de un conjunto de datos, así como para detectar valores atípicos y posibles sesgos en la distribución.**

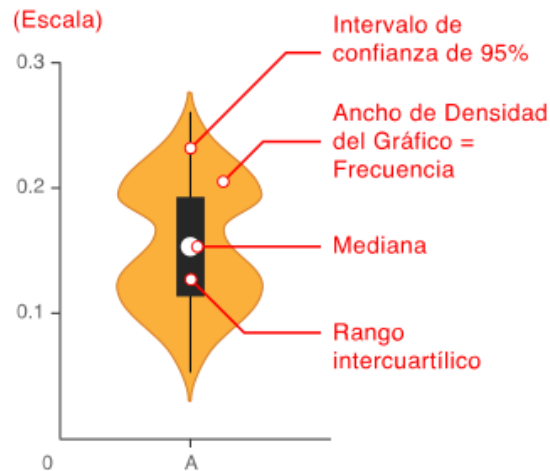


Definición: Diagrama de violín

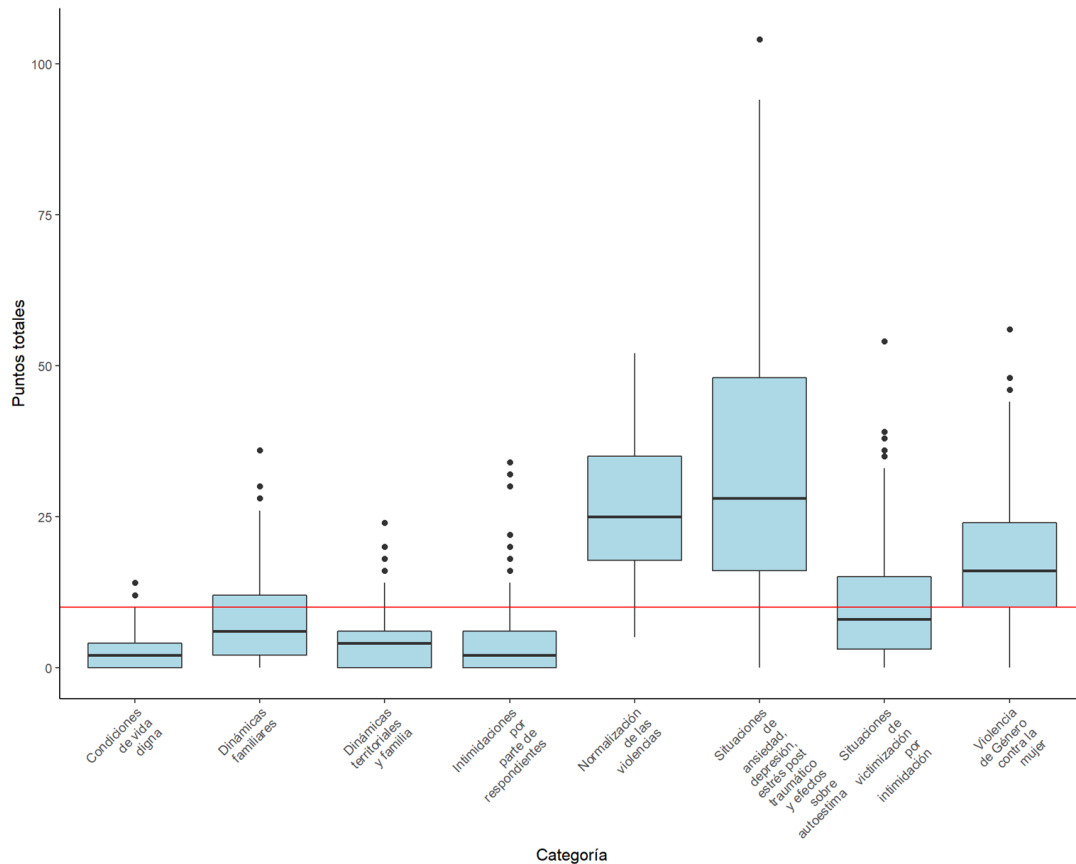
El gráfico de violín muestra la distribución de los datos mediante una forma de violín simétrica que se estrecha en los extremos y se ensancha en el centro. **La anchura del violín en cualquier punto representa la densidad de los datos en ese punto, y el tamaño del violín representa el número total de datos en ese conjunto.**

Al igual que con el gráfico de cajas y bigotes, se pueden mostrar puntos individuales para representar valores atípicos o extremos en los datos. Además, los gráficos de violín pueden mostrar varias distribuciones de datos en el mismo gráfico, lo que facilita la comparación entre grupos de datos.

El gráfico de violín **es útil para mostrar la forma, la simetría y la dispersión de los datos.**



Distribución de puntos por categoría



Definición de variable dependiente

Condiciones de asignación de nivel por categoría

Categoría	Condición de nivel
Condiciones de vida digna	Si puntos ≥ 4 , ALTO, si puntos ≤ 3 BAJO
Dinámicas familiares	Si puntos ≥ 4 , ALTO, si puntos ≤ 3 BAJO
Dinámicas territoriales y familia	Si puntos ≥ 4 , ALTO, si puntos ≤ 3 BAJO
Intimidaciones por parte de respondientes	Si puntos ≥ 10 , ALTO, si puntos ≤ 9 BAJO
Normalización de las violencias	Si puntos ≥ 10 , ALTO, si puntos ≤ 9 BAJO
Situaciones de ansiedad, depresión, estrés post traumático y efectos sobre autoestima	Si puntos ≥ 10 , ALTO, si puntos ≤ 9 BAJO
Situaciones de victimización por intimidación	Si puntos ≥ 10 , ALTO, si puntos ≤ 9 BAJO
Violencia de Género contra la mujer	Si puntos ≥ 10 , ALTO, si puntos ≤ 9 BAJO



Cuantificación por categoría y nivel

Categoría	Alto	% Alto	Bajo	% Bajo
Condiciones de vida digna	83	40%	125	60%
Dinámicas familiares	151	73%	57	27%
Dinámicas territoriales y familia	124	60%	84	40%
Intimidaciones por parte de respondientes	36	17%	172	83%
Normalización de las violencias	195	94%	13	6%
Situaciones de ansiedad, depresión, estrés post traumático y efectos sobre autoestima	182	88%	26	13%
Situaciones de victimización por intimidación	92	44%	116	56%
Violencia de Género contra la mujer	158	76%	50	24%
Violencia de Género contra la mujer (*solo mujeres)	79	40%	35	60%



Condiciones de asignación de nivel general

Para la asignación del nivel general se suman todos los puntos obtenidos por categoría y se clasifican según los siguientes criterios:

Género	Condición de nivel
Femenino	<ul style="list-style-type: none">• Si la puntuación obtenida es menor que 70, asignar nivel "BAJO"• Si la puntuación obtenida está entre 70 y 140, asignar nivel "MEDIO"• Si la puntuación obtenida es mayor o igual a 140, asignar nivel "ALTO"
Masculino	<ul style="list-style-type: none">• Si la puntuación obtenida es menor que 60, asignar nivel "BAJO"• Si la puntuación obtenida está entre 60 y 130, asignar nivel "MEDIO"• Si la puntuación obtenida es mayor o igual a 130, asignar nivel "ALTO" <p>NOTA: Se excluye del cálculo la categoría "Violencia de Género contra la mujer"</p>



Modelo

Primer aproximación

Variable dependiente:

Su situación de vulnerabilidad
es Alto, Medio o Bajo, no se
discriminación por categoría



Cuantificación por nivel

Nivel	Hombres	% hombres	Mujeres	%mujeres	# Estudiantes	%
Alto	29	32%	32	28%	61	30%
Medio	51	55%	58	51%	109	53%
Bajo	12	13%	24	21%	36	17%

*En este caso no se tienen en cuenta a los estudiantes LGTBI+



Ejemplo

De los 36 estudiantes que quedaron en Bajo se pudo evidenciar que

Número de categorías en alto	Número de estudiantes
1	4
2	13
3	11
4	7
5	1



Ejemplo

Podemos observar en el siguiente ejemplo:

ID Usuario	Nivel
247	Bajo

Categoría	Nivel
Condiciones de vida digna	BAJO
Dinámicas familiares	BAJO
Dinámicas territoriales y familia	BAJO
Intimidaciones por parte de respondientes	BAJO
Normalización de las violencias	ALTO
Situaciones de ansiedad, depresión, estrés post traumático y efectos sobre autoestima	ALTO
Situaciones de victimización por intimidación	BAJO
Violencia de Género contra la mujer	ALTO



Nueva propuesta

Modelo

Primer aproximación

Variable dependiente:

Su situación de vulnerabilidad es Alto, Medio o Bajo, no se discriminación por categoría

1 solo modelo

Segunda aproximación

Variable dependiente:

Su situación de vulnerabilidad es Alto, Medio o Bajo para tres categorías:

- Normalización de las violencias
- Situaciones de victimización por intimidación
- Situaciones de ansiedad, depresión, etc

3 modelos



La nueva clasificación depende del porcentaje de preguntas que respondieron en las diferentes etiquetas:

Nivel	Condición de nivel
BAJO	Si el 90% o más de las preguntas marca Nunca o si la suma de Nunca y Casi nunca suma 90% o más
MEDIO	Son los estudiantes que no tienen el nivel bajo ni alto
ALTO	Si Siempre es 33% o más, Casi siempre es 50% o más, suma de Casi siempre y siempre son más del 33%, o el



Niveles por categoría

Nivel	Alto	%alto	Medio	%medio	Bajo	%bajo
Normalización de las violencias	62	30%	97	47%	47	23%
Situaciones de victimización por intimidación	13	6%	34	16%	161	77%
Situaciones de ansiedad, depresión, etc	70	33%	53	26%	85	41%



Nivel	Alto	%alto	Medio	%medio	Bajo	%bajo
Intimidaciones por parte de respondientes	0	0%	7	3%	201	97%

