



GOBIERNO DEL ESTADO DE
VERACRUZ

SEGOB
SECRETARÍA DE GOBIERNO

COESPO
CONSEJO ESTATAL DE POBLACIÓN
DE VERACRUZ

DIAGNÓSTICO GEOESTADÍSTICO INTEGRAL DE RIESGOS Y VIOLENCIA ESCOLAR EN EL ESTADO DE VERACRUZ

Estrategias Basadas en Evidencia para la Prevención Social

REPORTE DE CONSULTORÍA ESTADÍSTICA

Elaborado por:

Leonardo Hernández Ávila
Miguel Ángel Ladrón de Guevara
César Alexis Ortiz

Contenido

INTRODUCCIÓN	2
RESULTADOS	4
Estadísticas descriptivas	4
Prueba de hipótesis.....	8
Supuesto de normalidad	8
Supuesto de homocedasticidad	13
ANOVA de Welch	15
Post-hoc Games-Howell.....	16
Conclusiones	18
Referencias.....	19

INTRODUCCIÓN

La Secretaría Técnica del Consejo Estatal de Población (COESPO) es un organismo descentrado de la Secretaría de Gobierno. A su cargo está el diseño y aplicación de la política de población en Veracruz. El reporte a continuación presenta los resultados del análisis de datos referentes a la Estrategia de Prevención Social de la Violencia y el Delito en Planteles Educativos del Estado de Veracruz a cargo de la anteriormente mencionada institución.

El objetivo del proyecto es desarrollar un análisis integral basado en la aplicación del instrumento (encuesta) aplicado a directivos de planteles educativos en el Estado de Veracruz; Identificando problemáticas, factores de riesgo y situaciones de violencia que afectan el entorno escolar. El fin último del proyecto es desarrollar políticas públicas que garanticen entornos seguros y favorables para el aprendizaje. Es imperativo establecer una estrategia que permita coordinar focalizar y ejecutar intervenciones oportunas entre dependencias gubernamentales.

Mencionado lo anterior, la implementación de técnicas estadísticas y procedimientos en el software correspondiente para rescatar información útil para los fines descritos es el enfoque principal. El instrumento analizado consta de 94 reactivos sin contar las etiquetas que identifican cada caso por medio del nombre de la institución, ubicación geográfica, municipio, localidad, responsable del plantel, entre otros. Las variables de interés fueron los índices de violencia y de discriminación y la etiqueta correspondiente al nivel educativo al que pertenece el caso (primaria, secundaria o media superior). El instrumento analizado consta de 94 reactivos, además de variables de identificación del plantel (institución, ubicación geográfica, municipio, localidad y responsable, entre otras). Las variables de interés en este reporte son los índices de violencia y de discriminación, así como el nivel educativo

(primaria, secundaria y media superior). Se cuenta con un total de 886 casos, seleccionados mediante un muestreo por conveniencia no probabilístico, por lo que no es posible generalizar los resultados a la población total de 23,032 planteles educativos del estado.

El análisis incluye técnicas descriptivas (media, mediana, varianza y desviación estándar) y el uso de diagramas de cajas y bigotes para comparar los niveles educativos. Adicionalmente, se contrasta la existencia de diferencias entre grupos mediante pruebas inferenciales: se evalúan los supuestos de normalidad, homocedasticidad e independencia de errores y, con base en sus resultados, se propone el uso de un ANOVA de Welch para identificar diferencias entre niveles educativos.

RESULTADOS

Estadísticas descriptivas

Tabla 1

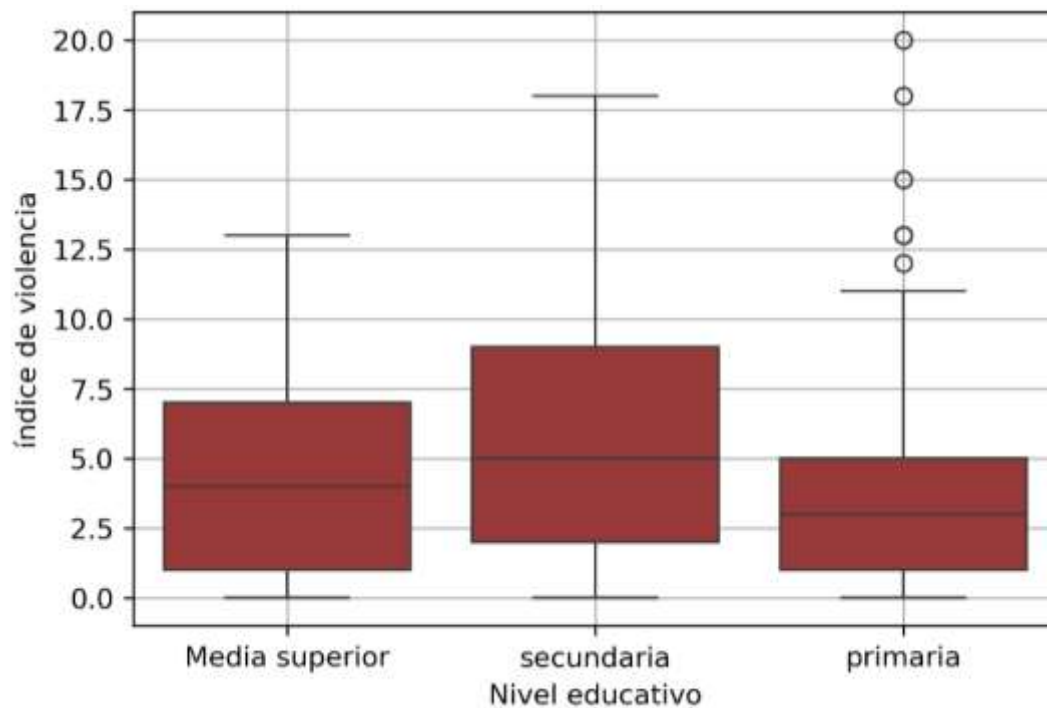
Estadísticas descriptivas del índice de violencia en función del nivel educativo

Nivel educativo	Número de muestra	Media	Mediana	Varianza	Desviación estándar
Primaria	319	3.64	3	10.47	3.24
Secundaria	332	5.19	5	15.71	3.96
Media superior	235	4.07	4	10.17	3.19

En las estadísticas descriptivas de la variable índice de violencia en función del nivel educativo (Tabla 1) de las escuelas encuestadas, se refuerza lo dicho con anterioridad, se detectan ciertos patrones relevantes, el nivel de secundaria presenta un promedio más alto de violencia a comparación de los otros grupos. Mientras que el nivel de primaria es el que presenta un promedio más bajo de violencia. Los 3 grupos están relativamente balanceados, lo que hace que esta comparación sea válida.

Figura 1

Diagrama de cajas y bigotes (boxplot) comparativo del índice de violencia por del nivel educativo



En relación con el índice de violencia, se detecta en el boxplot comparativo (Figura 1) que el nivel de secundaria presenta valores más altos a comparación de los otros niveles educativos, además, este grupo tiene una mayor variabilidad. Por otro lado, el nivel de primaria es el grupo que presenta valores más bajos, sin embargo, tiene algunos valores atípicos de ciertas escuelas en las que se detectó un nivel alto de violencia. Cabe recalcar que existe un traslape entre los tres gráficos, por lo que se tendría que aplicar una prueba de hipótesis para evaluar si sus medias son estadísticamente diferentes.

Tabla 2

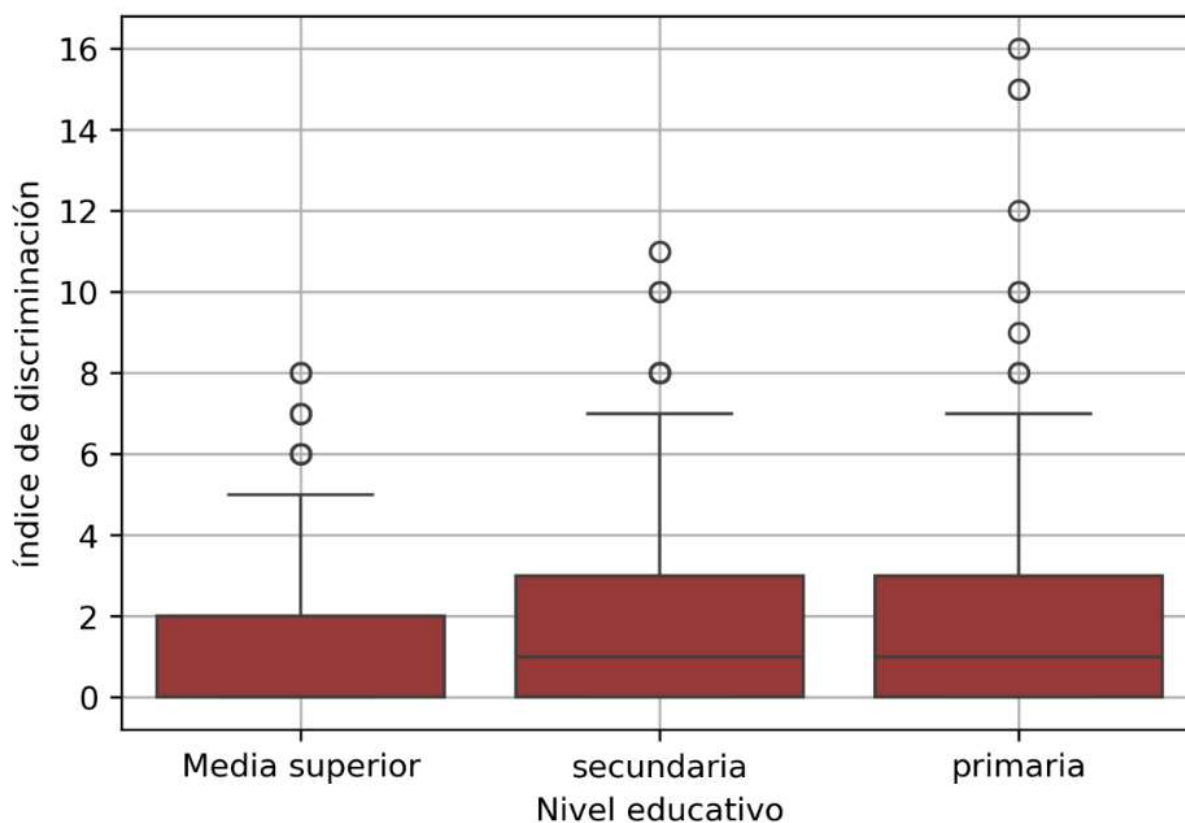
Estadísticas descriptivas del índice de discriminación en función del nivel educativo

Nivel educativo	Número de muestra	Media	Mediana	Varianza	Desviación estándar
Primaria	319	1.77	1	5.88	2.42
Secundaria	332	2.09	1	6.33	2.52
Media superior	235	1.34	0	3.94	1.98

En las estadísticas del índice de discriminación (Tabla 2) se aprecia que, para el grupo de media superior, el límite inferior coincide con la mediana (0). Además, los 3 grupos tienen valores muy bajos y debido a su baja variabilidad, los datos se concentran en niveles de discriminación bajos.

Figura 2

Boxplot comparativo del índice de discriminación en función del nivel educativo



Respecto al índice de discriminación, se observa en el gráfico (Figura 2) que los niveles de secundaria y de primaria parecen tener comportamientos muy parecidos, siendo el nivel de media superior el grupo con valores más bajos. Los tres grupos presentan valores atípicos, lo que puede representar un problema en un análisis de varianza clásico, por lo que está justificada la aplicación de otra prueba para evaluar la diferencia de medias.

Prueba de hipótesis

Con el objetivo de determinar si el nivel educativo se asocia con diferencias en el índice de violencia y de discriminación, se planteó una prueba de hipótesis para comparar los promedios de dicha variable entre los distintos niveles educativos. Dado que el índice de violencia y el índice de discriminación son variables cuantitativas y el nivel educativo es una variable cualitativa con tres categorías, la comparación de medias permite evaluar de manera formal si las diferencias observadas entre grupos pueden atribuirse al azar o si reflejan variaciones estadísticamente significativas en la población.

Variables:

$$X_1 = \text{índice de violencia}$$

$$X_2 = \text{índice de discriminación}$$

Supuesto de normalidad

Antes de realizar el análisis de varianza, es necesario evaluar los supuestos de normalidad y homocedasticidad. Dado el tamaño muestral elevado ($n = 886$), la prueba de Shapiro–Wilk resulta excesivamente sensible a desviaciones leves de la normalidad, mientras que la prueba de Kolmogorov–Smirnov presenta baja potencia. Por ello, se optó por la prueba de D’Agostino–Pearson, es la más adecuada para muestras grandes y enfocada en desviaciones relevantes (asimetría y curtosis), complementando el análisis con gráficos Q-Q.

Para la variable índice de violencia:

$$H_0: X_1 \text{ tiene una distribución normal}$$

$$H_1: X_1 \text{ no tiene una distribución normal}$$

Tabla 3

Prueba de D'Agostino-Pearson para evaluar la normalidad de la variable Índice de violencia en primaria

Estadístico	p-valor
$K^2 = 84.3462$	< 0.0001

La prueba de D'Agostino-Pearson (Tabla 3), con una confiabilidad del 95%, indica que existe suficiente evidencia estadística para asumir que el índice de violencia en primaria no se distribuye normalmente.

Tabla 4

Prueba de D'Agostino-Pearson para evaluar la normalidad de la variable Índice de violencia en secundaria

Estadístico	p-valor
$K^2 = 19.0805$	0.0001

Esta prueba dio como resultado que no existe normalidad en el índice de violencia en secundaria (Tabla 4).

Tabla 5

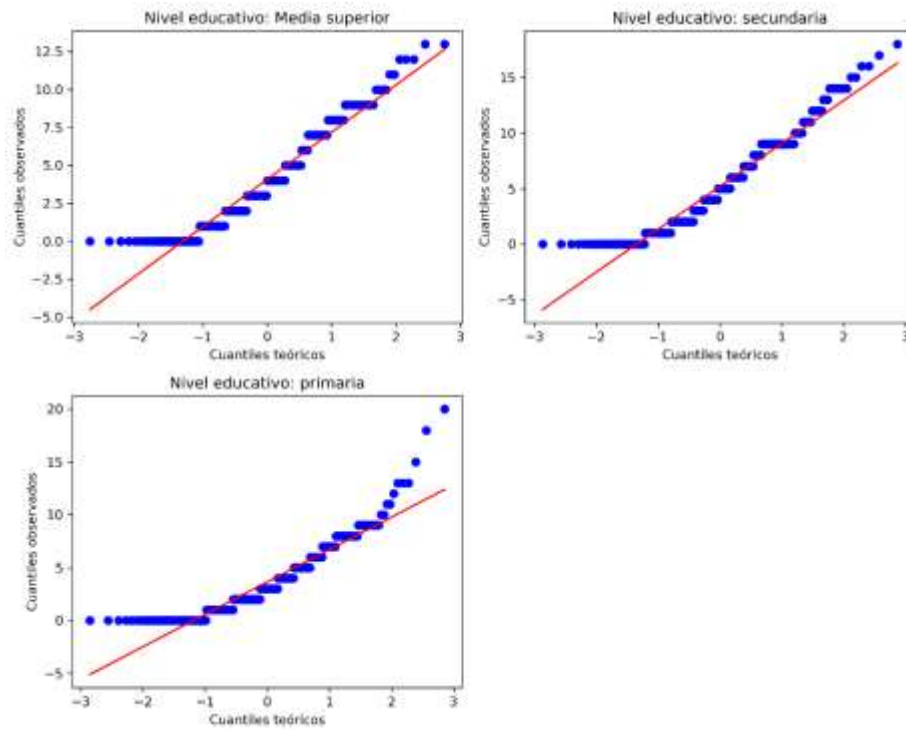
Prueba de D'Agostino-Pearson para evaluar la normalidad de la variable Índice de violencia en media superior

Estadístico	p-valor
$K^2 = 15.6450$	0.0004

La Tabla 5 muestra el rechazo de la hipótesis nula para el supuesto de normalidad, es decir, los datos del índice de violencia en nivel media superior no se distribuyen normalmente.

Figura 3

Gráficos Q-Q para evaluar la normalidad del índice de violencia por nivel educativo



Los gráficos Q-Q (Figura 3) confirman los resultados del supuesto de normalidad, los índices de violencia en función del nivel educativo no se distribuyen de forma normal.

Para la variable índice de discriminación:

$H_0: X_2$ tiene una distribución normal

$H_1: X_2$ no tiene una distribución normal

Tabla 6

Prueba de D'Agostino-Pearson para evaluar la normalidad de la variable Índice de discriminación en primaria

Estadístico	p-valor
$K^2 = 206.6037$	< 0.0001

En relación con el índice de discriminación, los resultados de la prueba de D'Agostino-Pearson (Tabla 6) indican que existe evidencia estadísticamente suficiente para asumir con un nivel de confianza del 95% que los datos en función del nivel de primaria no se distribuyen de forma normal.

Tabla 7

Prueba de D'Agostino-Pearson para evaluar la normalidad de la variable Índice de discriminación en secundaria

Estadístico	p-valor
$K^2 = 76.3837$	< 0.0001

La Tabla 7 muestra los resultados de la prueba de normalidad, la cual da como resultado, con un nivel de confianza del 95%, el rechazo de la hipótesis nula, es decir, el índice de discriminación en secundaria no se distribuye de manera normal.

Tabla 8

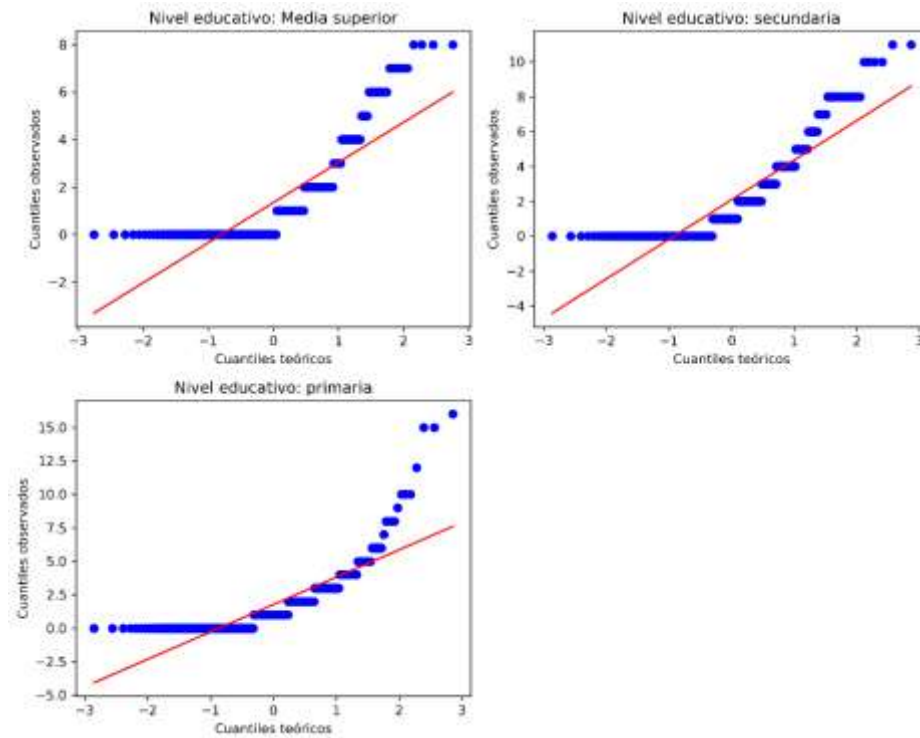
Prueba de D'Agostino-Pearson para evaluar la normalidad de la variable Índice de discriminación en media superior

Estadístico	p-valor
$K^2 = 78.0973$	< 0.0001

La Tabla 8 comprueba lo mismo para el índice de discriminación en el nivel media superior, los datos no se distribuyen de manera normal.

Figura 4

Gráficos Q-Q del índice de discriminación en función del nivel educativo



Los gráficos Q-Q confirman los resultados anteriores, el índice de discriminación en función del nivel educativo no se distribuye de forma normal.

Supuesto de homocedasticidad

Aunque la prueba de D'Agostino-Pearson indicó desviaciones de la normalidad, se evaluó la homogeneidad de varianzas mediante la prueba de Levene, contrastando de forma conjunta las varianzas del índice de violencia entre los tres niveles educativos, utilizando la mediana como medida de centralidad para robustez ante valores atípicos, ya que este supuesto es más relevante para la validez del ANOVA y su elección metodológica.

Con respecto a σ^2_k , donde $k = 1, 2$; representa las variables independientes:

σ^2_{k1} = La varianza poblacional de X_k para primaria

σ^2_{k2} = La varianza poblacional de X_k para secundaria

σ^2_{k3} = La varianza poblacional de X_k para media superior

Hipótesis Estadística:

$$H_0: \sigma^2_{k1} = \sigma^2_{k2} = \sigma^2_{k3}$$

$$H_1: \sigma^2_{ki} \neq \sigma^2_{kj}, \forall i \neq j; i, j=1,2,3$$

Para la variable Índice de violencia

$$H_0: \sigma^2_{X1,primaria} = \sigma^2_{X1,secundaria} = \sigma^2_{X1,media_superior}$$

$$H_1: \sigma^2_{X1,k_i} \neq \sigma^2_{X1,k_j}, \forall i \neq j; i, j=1,2,3$$

Tabla 9

Prueba de homocedasticidad (Levene) para la variable Índice de violencia

Estadístico	p-valor
W = 13.2089	< 0.001

La prueba de Levene (Tabla 9) para la variable índice de violencia, con un nivel de confiabilidad del 95% dio como resultado el rechazo de la hipótesis nula, es decir, al menos una de las varianzas de los tres niveles educativos es diferente.

Para la variable Índice de discriminación

$$H_0: \sigma^2_{X2,primaria} = \sigma^2_{X2,secundaria} = \sigma^2_{X2,media_superior}$$

$$H_1: \sigma^2_{X2,k_i} \neq \sigma^2_{X2,k_j}, \forall i \neq j; i, j=1,2,3$$

Tabla 10

Prueba de homocedasticidad (Levene) para la variable Índice de discriminación

Estadístico	p-valor
W = 4.6646	0.0097

La Tabla 10 muestra los resultados de la prueba de Levene para la variable índice de discriminación, la cual indica con una confiabilidad del 95% que existe heterogeneidad en los datos.

Con esto se concluye que no se cumplen los supuestos del ANOVA clásico. En consecuencia, se optó por aplicar un análisis de varianza de Welch, el cual es robusto ante varianzas desiguales y tamaños muestrales distintos entre los grupos, permitiendo una comparación válida de las medias del índice de violencia y de discriminación entre los niveles educativos. Además, respecto al supuesto de independencia, se consideró cumplido, dado que cada observación corresponde a un individuo distinto y no existen mediciones repetidas ni dependencia estructural entre las observaciones.

ANOVA de Welch

Con respecto a X_k , donde $k = 1, 2$; representa las variables independientes:

μ_{k1} = La media poblacional de X_k para primaria

μ_{k2} = La media poblacional de X_k para secundaria

μ_{k3} = La media poblacional de X_k para media superior

Hipótesis Estadísticas

$$H_0: \mu_{k1} = \mu_{k2} = \mu_{k3}$$

$$H_1: \mu_{ki} \neq \mu_{kj} \text{ , } \forall i \neq j; i, j = 1, 2, 3$$

Tabla 11

Análisis de varianza de Welch para el índice de violencia

Fuente	gl (numerador)	gl (denominador)	Estadístico F (Welch)	p-valor
Nivel educativo	2	567.94	15.3037	< 0.001

El análisis de varianza de Welch (Tabla 11) mostró diferencias estadísticamente significativas en el índice de violencia entre los niveles educativos ($F_{2, 567.94} = 15.30$, $p < 0.001$).

Tabla 12

Análisis de varianza de Welch para el índice de discriminación

Fuente	gl (numerador)	gl (denominador)	Estadístico F (Welch)	p-valor
Nivel educativo	2	579.41	7.8550	0.0004

El análisis de varianza de Welch (Tabla 12) mostró diferencias estadísticamente significativas en el índice de violencia entre los niveles educativos ($F_{2, 579.41} = 7.855$, $p < 0.001$).

Post-hoc Games-Howell

Tabla 13

Tabla de comparaciones post-hoc mediante el método de Games-Howell para el índice de violencia

Grupo A	Grupo B	Dif. de medias	Valor p	Tamaño del efecto
Media superior	Primaria	0.43	0.263	0.13
Media superior	Secundaria	-1.12	< 0.001	-0.31
Primaria	Secundaria	-1.55	< 0.001	-0.43

Las comparaciones post-hoc mediante el método de Games–Howell (Tabla 13) indicaron que el índice de violencia promedio en el nivel de secundaria es significativamente mayor que en primaria y media superior ($p < 0.001$). No se observaron diferencias estadísticamente significativas entre los niveles de primaria y media superior.

Tabla 14

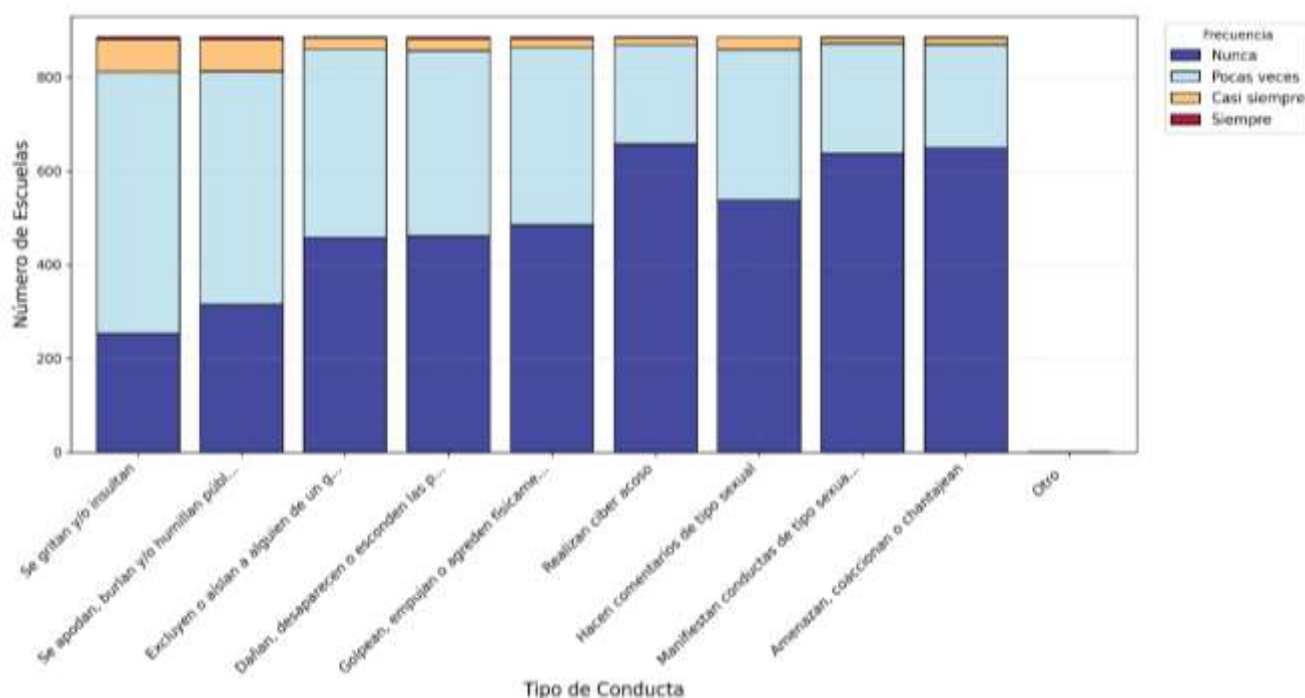
Tabla de comparaciones post-hoc mediante el método de Games-Howell para el índice de discriminación

Grupo A	Grupo B	Dif. de medias	Valor p	Tamaño del efecto
Media superior	Primaria	-0.42	0.062	-0.19
Media superior	Secundaria	-0.75	< 0.001	-0.32
Primaria	Secundaria	-0.32	0.219	-0.13

Las comparaciones post-hoc mediante el método de Games–Howell (Tabla 14) mostraron que el índice de discriminación promedio es significativamente mayor en el nivel educativo de secundaria en comparación con media superior ($p < 0.001$). No se observaron diferencias estadísticamente significativas entre primaria y media superior, ni entre primaria y secundaria. En general, los tamaños del efecto fueron pequeños, lo que indica que las diferencias encontradas, aunque relevantes, presentan una magnitud moderada.

Figura 5

Prevalencia de tipos de violencia en escuelas de Veracruz



El gráfico de barras apiladas (Figura 5) muestra que, en la mayoría de las escuelas, las distintas formas de violencia son reportadas principalmente en la categoría “Nunca”, lo que sugiere una baja prevalencia general. No obstante, existen diferencias importantes entre tipos de conducta. Las formas de violencia verbal y relacional, como los insultos, burlas y la exclusión, presentan una mayor proporción de respuestas “Pocas veces”, lo que indica que son las manifestaciones más comunes, aunque generalmente esporádicas.

Por otro lado, las conductas más graves, como la agresión física, las amenazas y las conductas de tipo sexual, son menos frecuentes en términos generales, pero cuando se presentan tienden a concentrarse en las categorías “Casi siempre” y “Siempre”, lo que sugiere que en algunos planteles estas prácticas son persistentes. El ciberacoso aparece con una presencia intermedia, lo que apunta a un fenómeno presente pero no generalizado. En conjunto, el gráfico evidencia una violencia escolar heterogénea, con distintos patrones según el tipo de conducta.

Conclusiones

El análisis estadístico permitió identificar diferencias claras en los índices de violencia y discriminación entre los niveles educativos de los planteles analizados en el estado de Veracruz. Los resultados descriptivos muestran que el nivel de secundaria concentra los valores promedio más altos tanto en el índice de violencia como en el índice de discriminación, lo que sugiere una mayor exposición de los estudiantes de este nivel a conductas y contextos de riesgo dentro del entorno escolar.

Las técnicas empleadas permitieron comparar de manera rigurosa las diferencias entre grupos. La evaluación de los resultados de las pruebas de los supuestos estadísticos fue evidencia para justificar la elección del ANOVA de Welch como la prueba más adecuada, dada la heterogeneidad observada en las varianzas.

Al tratarse de un muestreo no probabilístico, no es posible generalizar los resultados al total de planteles educativos del estado; sin embargo, los hallazgos proporcionan evidencia valiosa y orientativa para el diseño de estrategias de prevención de la violencia escolar. Teniendo en cuenta lo anterior, se sugiere que los esfuerzos institucionales de la COESPO y otras dependencias involucradas enfatizen la intervención en planteles de nivel secundaria.

Para enriquecer y facilitar la comprensión de los resultados del análisis, se generaron mapas con los índices de discriminación y violencia, así como mapas de calor, el resultado se puede consultar en el siguiente enlace: <https://leonardoavila2004.github.io/Mapas-Veracruz/>

En conjunto, los resultados evidencian la importancia del análisis estadístico aplicado a políticas públicas, ya que permite transformar datos empíricos en información útil para la toma de decisiones, la focalización de recursos y en este caso concreto la planificación de estrategias de prevención de violencia en los centros educativos de Veracruz.

Referencias

- Montgomery, D. C. (2019). *Design and Analysis of Experiments* (10th ed.).
- Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38(3–4), 330–336.
- Levene, H. (1960). Robust tests for equality of variances. En I. Olkin (Ed.), *Contributions to Probability and Statistics* (pp. 278–292). Stanford University Press.
- Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346), 364–367.
- Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, 37(3–4), 409–428.
- Kish, L. (1965). *Survey Sampling*. John Wiley & Sons.