



CINDE: Resultados

— Viernes 22 de abril —

Contenido

- Objetivo
- Datos
- Procesamiento y limpieza
- Variables dependientes
- Variables independientes
- Modelo
- Resultados de modelos



Objetivo

Objetivo

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 .

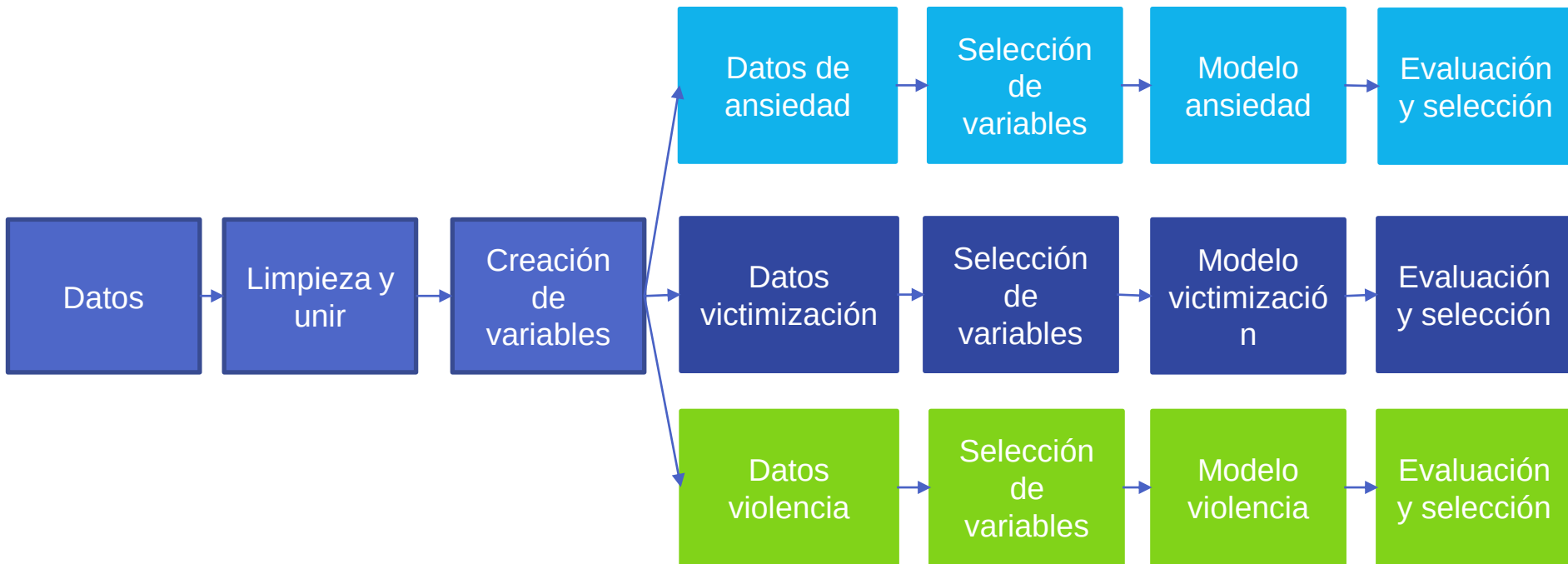
Específicamente en situación de vulnerabilidad por:

- Estar en una situación de **victimización por intimidación**
- Estar en una situación de **normalización de violencia**
- Estar en una situación de **depresión, ansiedad, estrés post traumático y efectos sobre autoestima**



Proceso

Proceso



Datos

Observaciones iniciales

- Caracterización de jóvenes:
 - 538 jóvenes
- Instrumento:
 - 447 jóvenes
- Comentarios
 - Lección con más jóvenes : 262 jóvenes (Lección 04 Caperucita roja)
 - Lección con menos jóvenes: 126 jóvenes (Lección 01 Bienvenida y presentación)

Nota: Fecha de corte 15 de abril de 2023



Procesamiento y limpieza

Procesamiento y limpieza : Datos de caracterización

Como parte del proceso de procesamiento y limpieza se realizaron los siguientes pasos:

- Retirar jóvenes por fuera del rango de 12 a 18
- Retirar jóvenes sin sexo
- Retirar jóvenes sin estrato



Procesamiento y limpieza: Datos instrumento

Como parte del proceso de procesamiento y limpieza se realizaron los siguientes pasos:

- Eliminar individuos que no fuera estudiantes
- Eliminar individuos que no hubieran respondido todas las preguntas del cuestionario
- Eliminar individuos que no tuvieran información completa en la caracterización



Procesamiento y limpieza: Datos comentarios

Como parte del proceso de procesamiento y limpieza de los comentarios, se realizaron los siguientes pasos:

- Dejar lecciones 132, 135, 142,143,144 y 145 (tienen más jóvenes)
- Eliminar comentarios duplicados
- Unificar comentarios por usuario y lección
- Eliminar números, puntuación y símbolos
- Pasar todo a minúsculas
- Eliminar tildes



Procesamiento y limpieza: Datos comentarios

Como parte del proceso de procesamiento y limpieza de los comentarios, se realizaron los siguientes pasos:

- Eliminar stopwords, palabras que no aportan ningún significado (e.j. “yo”, “eso”, “la”, “cual”, “más”, “que”)
- Eliminar patrones de letras (e.j. “aaaaa...”, “jajaja...”, “lalala...”)
- Eliminar palabras relacionadas con las preguntas (e.j. nombres de canciones)
- Lematización (e.j. “cuidado”, “cuidada”, “cuidadoso”, “cuidadosa” = “cuidar”)



Procesamiento y limpieza: Datos comentarios

Ejemplo

Antes	Después
<p>1.yo me cuido siempre estando acompañado por un adulto o persona que nos pueda cuidar,si es el caso de estar solos siempre tener un instrumento de proteccion como:gas pimienta,tener en el celular previamente marcado 123.</p> <p>2.recomiendo a mi familia y a mi escuela que para cuidarnos deberian tener camaras de seguridad o siempre tener un docente en los lugares mas remotos de la institucion,para la familia es estar con todo momento en contacto con ellos.</p>	<p>siempre ir acompañado adulto poder cuidar casos soledad siempre tener instrumento proteccion gas pimienta tener celular previamente marcado tener camara seguridad siempre contacto</p> <p>recomendar familia colegio cuidar deber tener camara seguridad siempre tener docente lugares remoto colegio familia momentos contacto</p>



Observaciones finales

- Caracterización de jóvenes:
 - Antes de limpieza: 538 jóvenes
 - Después de limpieza: 484 jóvenes con información completa
- Instrumento:
 - Antes de limpieza: 447 jóvenes
 - Después de limpieza inicial: 350 jóvenes
- Unión de caracterización e instrumentos
 - 306 jóvenes
- Unión de caracterización, instrumentos y comentarios:
 - **274 jóvenes**



Lecciones seleccionadas

Lección	Pregunta	# Observaciones
Caperucita roja	¿Qué estrategias se nos ocurren para cuidarnos y cuidar a otras niñas, niños y jóvenes, ante posibles violencias? ¿Qué mensaje le darías a la escuela y a la familia para fortalecer los entornos protectores?	262
Me llegó carta	Luego de leer la carta de Charles Chaplin, resalta las palabras que más te gustaron relacionadas con el cuidado y el amor propio...	260
Seis claves	Te invitamos a responder las claves para la reflexión, de acuerdo con el infografico. 1.¿Cómo podemos identificar riesgos y violencias para tomar medidas de protección?...	248
Reconociendo otras vivencias	Selecciona una imagen de la galería y apartir de ella cuéntanos una anécdota sobre situaciones similares que hayas vivido o visto en tu contexto escolar...	242
Reflexionemos nuestro entorno	Escucha las canciones que te hemos propuesto y sugiere, a partir de tus reflexiones, el material que conoces para la prevención de las violencias y la movilización ante la misma en los diferentes entornos.	236
Introducción	¿Qué tan cómodo me siento en espacios como la familia y la escuela conversando sobre mi identidad de género?	231



Variables del modelo

Modelo y cálculo de variables dependientes

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

Nivel	Ansiedad	Victimización	Violencia
BAJO	Si el 90% o más de las preguntas marca Nunca o si la suma de Nunca y Casi nunca suma 90% o más	Si el 90% o más de las preguntas marca Nunca, o si la suma de Nunca y Casi nunca suma 83% o más, y Siempre no tiene ninguna marca	Si el 90% o más de las preguntas marca Nunca, o si la suma de Nunca y Casi nunca suma 87% o más
ALTO	Son los estudiantes que no tienen el nivel bajo		



Variable dependiente - nivel_ansiedad

Posterior al cálculo de la variable dependiente nivel_ansiedad, se obtuvo la siguiente distribución por categoría:

Nivel	# Observaciones	Frec. relativa
BAJO	118	56.9%
ALTO	156	43.1%



Variable dependiente - nivel_victimizacion

Posterior al cálculo de la variable dependiente nivel_victimizacion, se obtuvo la siguiente distribución por categoría:

Nivel	# Observaciones	Frec. relativa
BAJO	225	82.1%
ALTO	49	17.9%



Variable dependiente - nivel_violencia

Posterior al cálculo de la variable dependiente nivel_violencia, se obtuvo la siguiente distribución por categoría:

Nivel	# Observaciones	Frec. relativa
BAJO	110	59.6%
ALTO	162	40.4%



Variables independientes

Se trabaja con más de **400** variables independientes.

Las variables sociodemográficas utilizadas en el modelo son:

Variable	Valores
estrato	numérico entre 1 y 5
etnia	"N" o "S", dependiendo si pertenece o no a un grupo étnico
edad	numérico entre 12 y 18
sexo	"m", "h", "otro"



Variables independientes

Características de los comentarios

Variable	Descripción
n_caracteres	número de caracteres en el comentario
n_palabras	número de palabras en el comentario
dif_nchar_median	desviación en el número de caracteres respecto a la mediana del número de caracteres en el comentario de los jóvenes con un nivel bajo por lección
dif_npal_median	desviación en el número de palabras respecto a la mediana del número de palabras en el comentario de los jóvenes con un nivel bajo



Variables independientes

Presencia del top 10 de palabras con mayor presencia en lecciones: numérico entre 0 y 1, 1 indica que el joven utilizó la palabra en su comentario y 0 indica que no

Palabras lección 132		
mismo	bien	feliz
tristeza	cosas	ser
gustar	querer	vida
amistad		



Variables independientes

Frecuencia de términos – Frecuencia inversa del document (TF-IDF) para palabras comunes entre todos los jóvenes.

Cuanto mayor sea la puntuación, más relevante será esa palabra para ese individuo en particular.



Variables independientes

Presencia de parejas de palabras más comunes en las lecciones:

Ejemplo	Valores
sentir comodidad	numérico entre 0 y 1, 1 indica que el joven utilizó la pareja en su comentario y 0 indica que no



Variables independientes

Presencia de palabras alerta

Tipo	Palabras
Ansiedad	"tristeza", "psicologo", "triste", "solo", "soledad", "desesperanza", "desmotivacion", "agotamiento", "irritabilidad", "insomnio", "pesimismo", "culpa", "estres", "mortificar", "mortificado", "desinteres", "desesperacion", "preocupacion", "negativo", "inutil", "no resaltar", "fatiga", "desinteres", "aislamiento", "suicidio", "humor", "verguenza", "no agradecer", "agujero", "dificultad", "suicida", "dificultad", "sin esperanza", "inseguridad", "desaparecer", "no escapar", "estresado", "incomodo", "miedo", "molesten", "criticas", "bullying", "temor", "angustia", "bulin", "morir", "trauma", "bullyg", "bullyn", "inseguro", "denigracion", "angustia", "dolio", "solo", "morir", "traumas", "mundo caer", "apodos", "sin autoestima", "baja autoestima", "ansiedad"
Violencia	"abuso", "bullying", "maltrato", "acoso", "violacion", "agresion", "amexas", "coaccion", "control", "dominacion", "marginacion", "humillacion", "exclusion", "rechazo", "discriminacion", "racismo", "xenofobia", "homofobia", "intimidacion", "chantaje", "manipulacion", "aislamiento", "ignorada", "burlas", "gritos", "ofensas", "insultos", "agresiones", "violencia", "temor", "acosan", "maltrato", "angustia", "dolio", "doler", "miedo", "vulnerable", "traumas"



Modelos

Modelos utilizados

Para llevar a cabo este análisis, se utilizaron cuatros modelos de aprendizaje automático diferentes:

- Random Forest
- Naive Bayes
- Stochastic Gradient Boosting
- Elastic - Net Regularized Generalized Linear Model

Se dividió el conjunto de datos en dos partes: el conjunto de entrenamiento y el conjunto de prueba. Se utilizó un enfoque de partición 80/20, donde el 80% de los datos se usaron para entrenar los modelos y el 20% restante se usó para evaluar su desempeño.



Modelo: Ansiedad

Matriz de confusión - Nivel de ansiedad

Modelo: Elastic Net

		Observación	
		Alto	Bajo
Predicción	Alto	27	22
	Bajo	4	11

		Observación	
		Alto	Bajo
Predicción	Alto	Verdaderos Positivos	Falsos Positivos
	Bajo	Falsos Negativos	Verdaderos Negativos



Matriz de confusión - Nivel de ansiedad

Modelo: Elastic Net

		Observación	
		Alto	Bajo
Predicción	Alto	27	12
	Bajo	4	11

	Alto
Exactitud	$70.37\% = (27 + 11) / 54$
Sensibilidad	$87.10\% = 27 / (27 + 4)$
Especificidad	$47.83\% = 11 / (11 + 12)$
Precisión	$69.23\% = 27 / (27 + 12)$



Selección del modelo de ansiedad

Métrica	Bosque Aleatorio	NB	SGBboosting	ElasticNet
Exactitud	66.67%	62.96%	66.67%	70.37%
Precisión	65.12%	62.79%	69.70%	69.23%
Especificidad	34.78%	30.43%	56.52%	47.83%
Sensibilidad	90.32%	87.10%	74.19%	87.10%



Modelo seleccionado para ansiedad: Elastic Net

- Sociodemográficas: sexo
- Presencia de palabras: confianza, riesgo, familia
- Frecuencia de palabras asociadas a sentimientos: tristeza y rabia
- Frecuencia de palabra: angustia, miedo
- Presencia de palabras bigramas: no hablar, no sentir, familia no, no poder, no gustar, no tener



Ejemplo

Usuario 594: Real: Alto / Predicción : Alto

Lección	Comentario
03 Me llegó carta	<ul style="list-style-type: none">- recuerdo feliz : cuando comienzo a dibujar lo que me gusta y cuando lo termino me da mucha felicidad ya que puedo ver como voy progresando cada vez mejor- recuerdo triste: <i>casi todo el tiempo en el colegio me trataron horrible con mucho bullying y la mayor parte de mi juventud a estado muy deprimido</i>- me esfuerzo mucho con las cosas que me gusta y no paro hasta terminarlo y que sea mejor que las cosas que he hecho antes- hago cosas que me agradan para sentirme bien y también paso tiempo con personas que me tratan bien como dibujar, jugar y pasar tiempo con mis amigos
02 Reconociendo otras vivencias	a pandemia fue lo mejor para mi ya que soy <i>una persona que no me gusta relacionarse con la gente o de por si salir a fuera</i> puede sentirme super bien tranquilo en mi casa en un lugar seguro y las clases virtuales lo mejor para ya que no conocia a nadie y mis amigos en esos tiempo también se volvieron desconocidos para mi y pude hacer mis trabajos a mi manera y una manera tranquila y constante no tenia a nadie como amigo como en virtualidad o en presencialidad y no sufrí esa falta de comunicación ya que me gusta estar solo quiero volver a esos tiempos



Ejemplo

Usuario 291: Real: Bajo / Predicción : Bajo

Lección	Comentario
03 Me llegó carta	1.R:honestidad,honradez,tolerancia,comprensión,respeto,solidaridad. 2.R:cuando me regalaron un computador sin importar todas las dificultades. 3.R:me siento estupendo ya que comparto una ambiente agradable con las personas que me rodean(colegio),me siento integrado en todo momento y no me desprecian. soy la mejor persona ya que mi compañerismo es enorme y rompe todas las barreras. 4.R:si 5.R:si
02 Reconociendo otras vivencias	No cuenta con datos



Ejemplo

Usuario 386: Real: Alto / Predicción : Bajo

Lección	Comentario
03 Me llegó carta	Juan José me escribo esta carta a mi para que puedas lograr todo lo que te propongas, siéntete muy orgulloso de la persona que eres, tu empatía tu confianza lo respetuoso que eres, nunca dejes de ser una persona tan feliz y positiva. \Aprecio las personas que están a mi alrededor que siempre me acompañan, mis amigos y mis familiares, el tener siempre alguien que me apoye hasta en los peores momentos. \r\nUn momento sumamente triste fue la muerte de mi abuela la cual paso conmigo muchos momentos de mi vida desde que tengo conciencia. Los momentos mas felices que he pasado son las reuniones familiares ya que mi familia es muy agradable. Me siento una persona muy confiada y segura de lo que podrá lograr en un futuro. Te aconsejo que nunca pierdas ese optimismo y carisma
02 Reconociendo otras vivencias	Experiencias en la pandemia: A mi familia en marzo de 2021 nos contagiamos de covid, al principio solo fue mi papa, pero ya días despues ya todos estabamos enfermos, mi mama yo y mi hermana, mis papa en especifico estuvo en una grave situación ya que el sufre de varias enfermedades, tuvimos que aislarnos por mucho tiempo y utilizar tapabocas en cada momento, gracias a Dios superamos ese momento sin ninguna perdida y pudimos sguir adelante con nuestras vidas sin ningun tipo de repercusion. Pero desde hay para adelante hemos tomado muchas medidas de precaución.



Modelo: Victimización

Matriz de confusión - Nivel de victimización

Modelo seleccionado Elastic Net

		Observación	
		Alto	Bajo
Predicción	Alto	7	7
	Bajo	2	36



Matriz de confusión - Nivel de victimización

Métrica	Bosque Aleatorio	NB	SGBosting	ElasticNet
Exactitud	80.77%	84.62%	82.69%	82.69%
Precisión	0%	57.14%	NA	50%
Especificidad	97.67%	93.02%	100%	83.72%
Sensibilidad	0%	44.44%	0%	77.78%



Modelo seleccionado para victimización: Elastic Net

- Sociodemográficas: sexo, edad
- Presencia de palabras: confianza, hacer, poder,
- Frecuencia de palabras asociadas a sentimientos: tristeza y rabia
- Frecuencia de palabra: soledad, comodidad
- Diferencia en el número de caracteres respecto a la mediana de ansiedad, victimización y violencia para la lección 144
- Presencia de palabras bigramas: entorno_familia, no querer, no hacer
- IDF: mensajes, respecto, alguien, siento, violencia, acoso, sexo



Ejemplo

Usuario 305: Real: Alto / Predicción : Alto

Lección	Comentario
03 Me llegó carta	1. la honestidad, empatía, solidaridad, amor, compromiso, justicia, integridad, verdad, colaboración, alegría. \r\n2. pasar tiempo en familia y rescatar animalitos.\r\n3. siendo sincera me siento cómoda en cualquier entorno, me siento bien conmigo y con los demás.
02 Reconociendo otras vivencias	no todos la pasamos bien\" me llamo la atención por que en algún momento de mi vida me sentí igual



Ejemplo

Usuario 312: Real: bajo / Predicción : bajo

Lección	Comentario
03 Me llegó carta	No la respondió
02 Reconociendo otras vivencias	me gusto mucho la primera imagen porque es completamente cierto, no todos nos entimos bien, no siempre estamos en aptas condiciones y eso es algo que es completamente normal y natural



Ejemplo

Usuario 291: Real: Alto / Predicción : Bajo

Lección	Comentario
03 Me llegó carta	mis momentos mas felices fueron cuando empecé a patinar a mis 6 años es lo que ,mas me gusta y me libera de mi rabia de mi tristeza me siento libre \n\n\ncuando empecé a bailar a mis 10 años me ayudo hacer las cosas con fuerza constancia rapidez
02 Reconociendo otras vivencias	a salud mental es uy importante ya que ha veces os estresamos por que no sabemos hacer las cosas o tenemos problemas en en el colegio o la casa entonces eso nos afecta mucho la mente



Modelo: Violencia

Matriz de confusión - Nivel de violencia

Modelo seleccionado : SGBoosting

		Observación	
		Alto	Bajo
Predicción	Alto	29	7
	Bajo	3	15



Matriz de confusión - Nivel de violencia

Modelo seleccionado : SGBosting

Métrica	Bosque Aleatorio	NB	SGBosting	ElasticNet
Exactitud	66.67%	59.26%	81.48%	74.07%
Precisión	71.88%	60.87%	80.56%	73.68%
Especificidad	59.09%	18.18%	68.18%	54.55%
Sensibilidad	71.88%	87.5%	90.62%	87.5%



Modelo seleccionado para violencia: SGBoosting

- Sociodemográficas: sexo, estrato
- Presencia de palabras: tener, familia, ser, malo, confianza, hogar
- Frecuencia de palabras asociadas a sentimientos: rabia y negatividad
- Frecuencia de palabra: comodidas, miedo
- Diferencia en el número de caracteres respecto a la mediana de violencia para la lección 143
- IDF: sentir
- Presencia de palabras bigramas: sentir comodidas, no poder, no tener



Ejemplo

Usuario 294: Real: Alto / Predicción : Alto

Lección	Comentario
03 Me llegó carta	Reconoce los valores que tienes. \r\nRecuerda los momentos más significativos de tu vida: de alegría, felicidad o tristeza.
02 Reconociendo otras vivencias	cuando me siento seguro y con q personas me siento seguro en los espacios abiertos espacios libres



37. Permanezco callado ante una situación de agresión, acoso e intimidación	Siempre
38. Intervengo cuando hay una situación de agresión, acoso e intimidación	Casi nunca
39. Denuncio ante adultos una situación de agresión, acoso e intimidación	Nunca
40. Permito la realización de juegos agresivos	Nunca
41. Veo acciones contundentes ante una situación de agresión, acoso e intimidación	Nunca
42. Los adultos tienen derecho a corregirme física y verbalmente por mi mal comportamiento	Casi Siempre
43. Por ser hombre debo ser tolerante ante situaciones de agresión, acoso e intimidación	Casi Siempre
44. Mi relación con los otros está mediada por palabras groseras y agresivas	Nunca
45. Los otros tienen razones válidas para agredirme	Nunca
46. Recibir golpes, insultos, comentarios agresivos o exclusiones es algo común en la escuela	Nunca
47. Los adultos pasan por alto las situaciones de agresión, acoso e intimidación	Nunca
48. Los adultos minimizan las denuncias por agresión, acoso e intimidación	Nunca



Ejemplo

Usuario 293: Real: bajo / Predicción : bajo

Lección	Comentario
03 Me llegó carta	Yo María Camila soy una persona maravillosa y valiente , estoy muy orgullosa de mi misma. me siento muy orgullosa de todo lo he logrado y me he propuesto. Me amo a mi misma, soy una personas muy amable, me gusta ayudar a las personas que lo necesitan, me gusta la música, me gusta pintar y hacer deportes. soy muy sentimental, no me gusta que me molesten y me han cosas que no me gustan.
02 Reconociendo otras vivencias	en mi familia no existe tipos de violencia. pero he visto mucha violencia en mi entorno donde vivo. \r\n mi recomendación seria que no nos quedemos callados o calladas sobre algo que estamos viviendo o pansado en nuestra familia y en el colegio, demos actuar para así hacer algo al respecto de lo que nos esta ocurriendo o sucediendo



Ejemplo

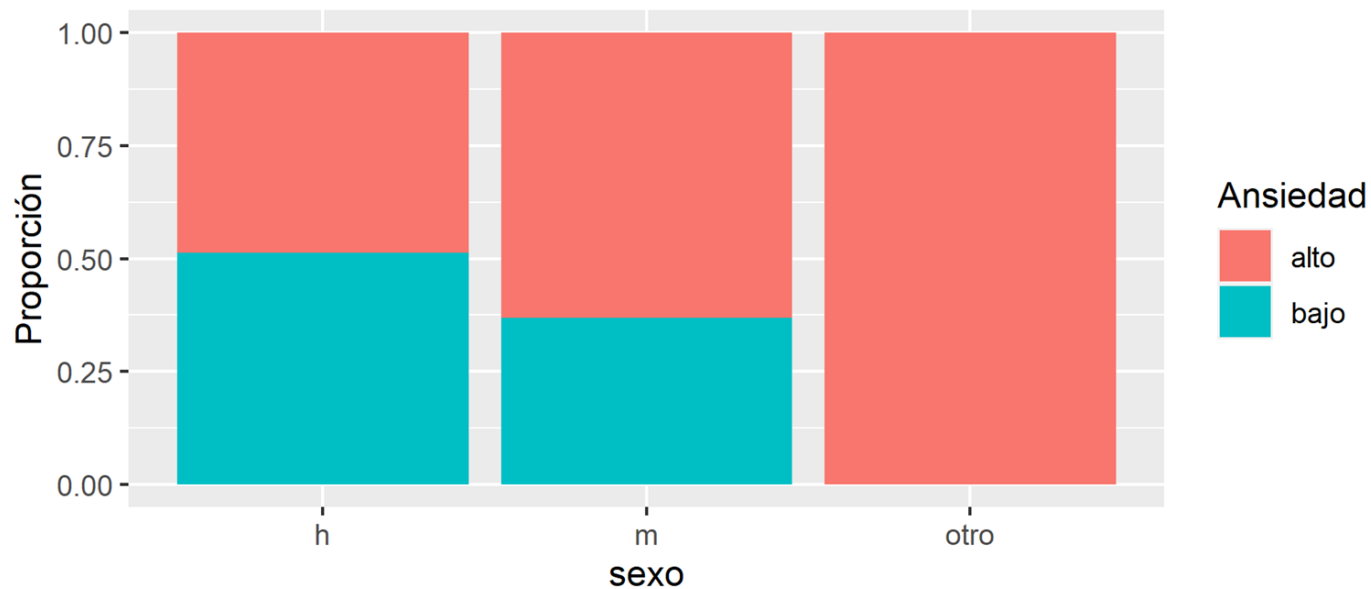
Usuario 383: Real: Alto / Predicción : Bajo

Lección	Comentario
03 Me llegó carta	Me siento tranquila, aunque a veces tengo muchas preocupaciones, en la mayoría de problemas se salir adelante sola o con ayuda, aunque no puedo negar que algunos problemas a los que me enfrento son difíciles o lo son para mi porque no tengo las capacidades suficientes para enfrentarlo o simplemente porque creo que es un obstáculo tan grande para mi que no seria capaz de enfrentarlo así que no intento cambiar. Pero no siempre, creo que he enfrentado obstáculos muy grandes para mi como para no seguir intentándolo para ser mejor cada día y poder ser lo que me proponga, no son muchas las veces en las que me pongo a pensar en mis capacidades: respetuosa, amable, soy una persona muy callada que no tiene demasiadas ganas de socializar, sea porque es un gran reto para mi como por ser capaz porque muchas veces \"no\" me lo permito
02 Reconociendo otras vivencias	Estaba en una Reunión virtual cuando derrepente se une a la Reunión una persona con un nombre ofensivo y empieza a poner las cosas fuera del contexto de la Reunión, por suerte no hubo ninguna situación que terminara en agresiones y se pudo solucionar expulsándolo de la Reunión

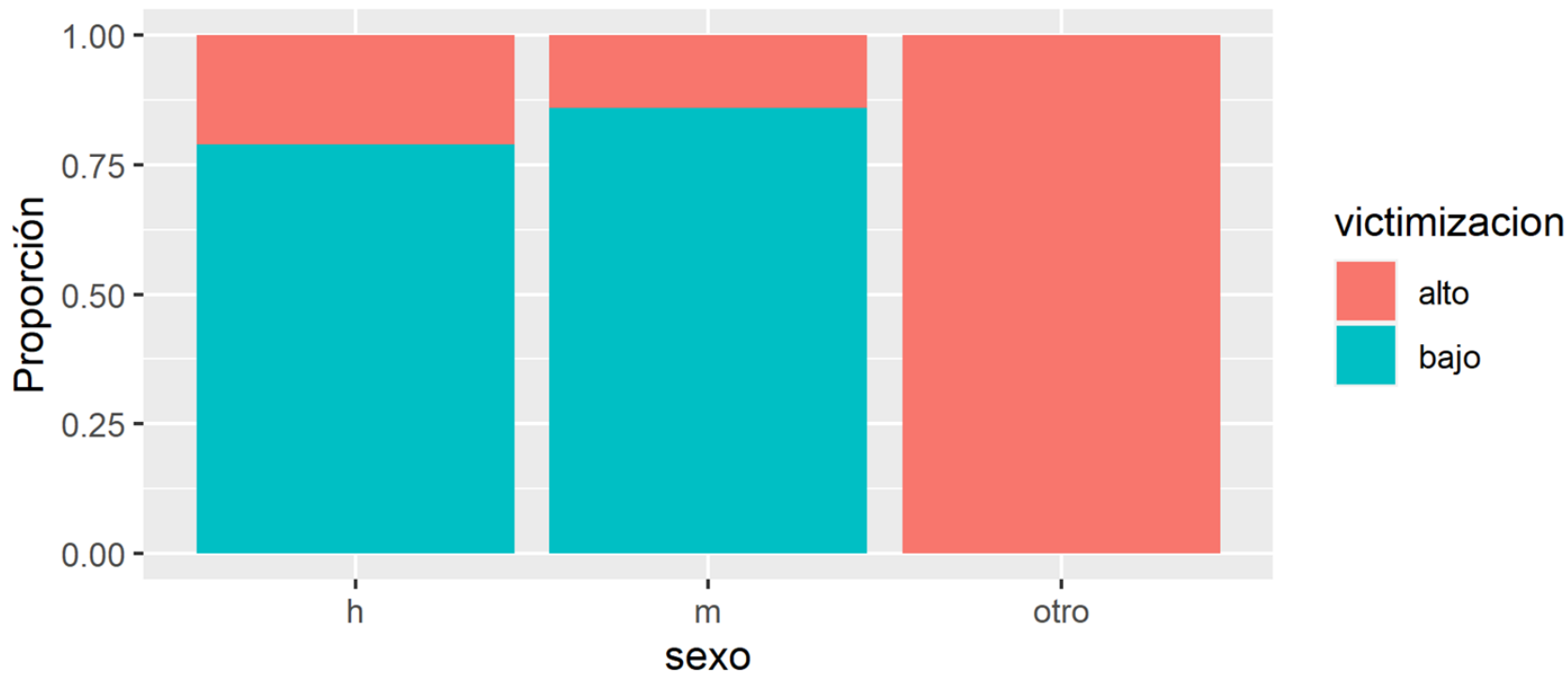


Resultados

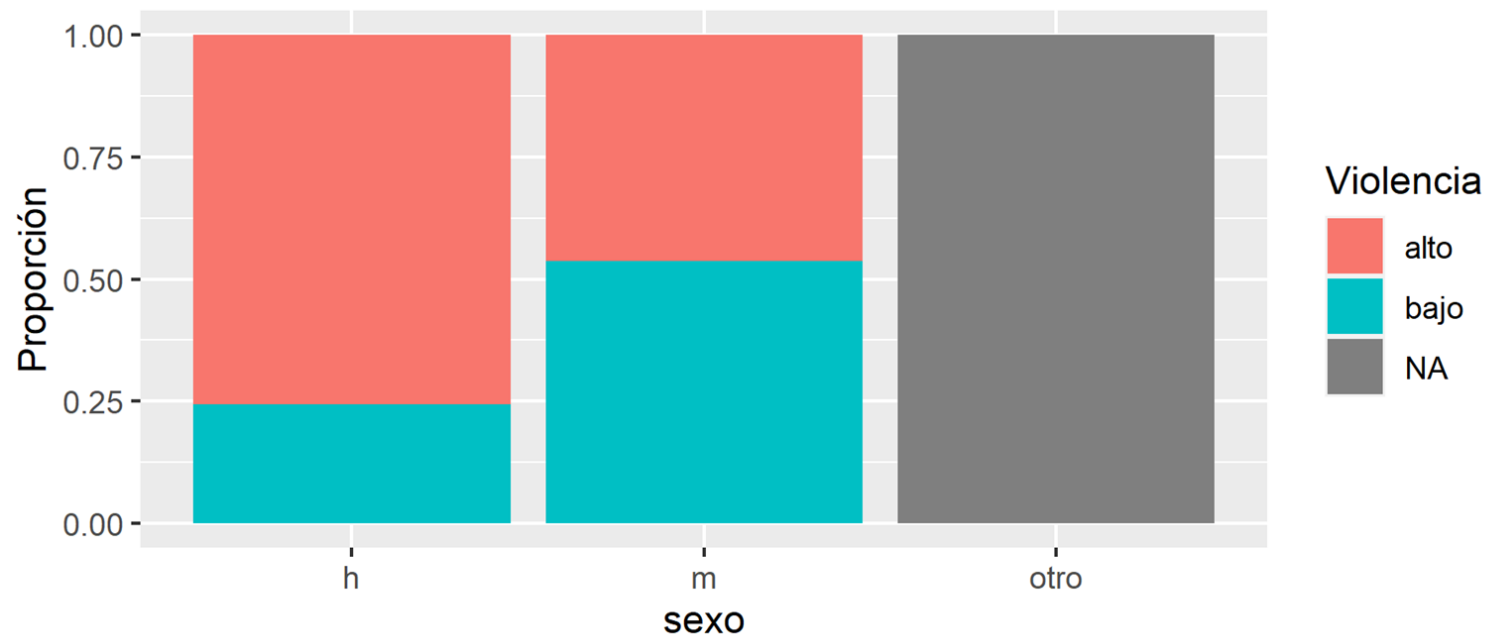
Niveles de ansiedad por sexo



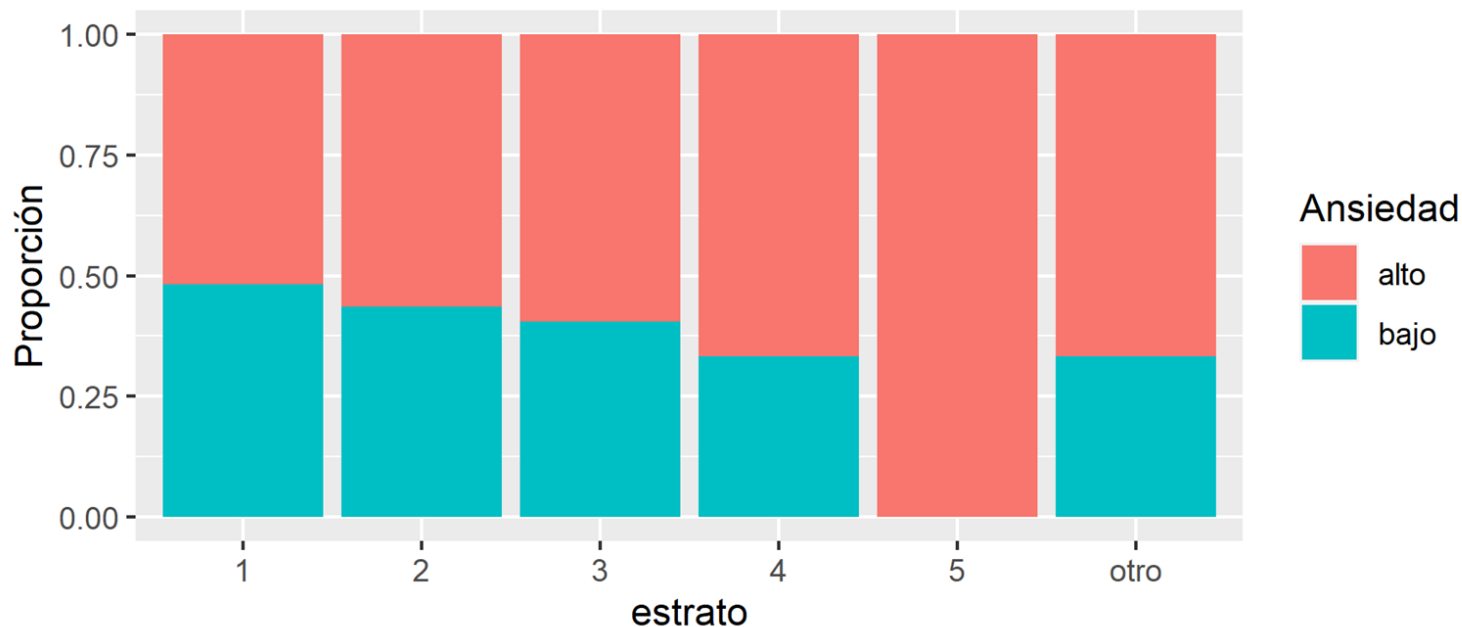
Niveles de victimización por sexo



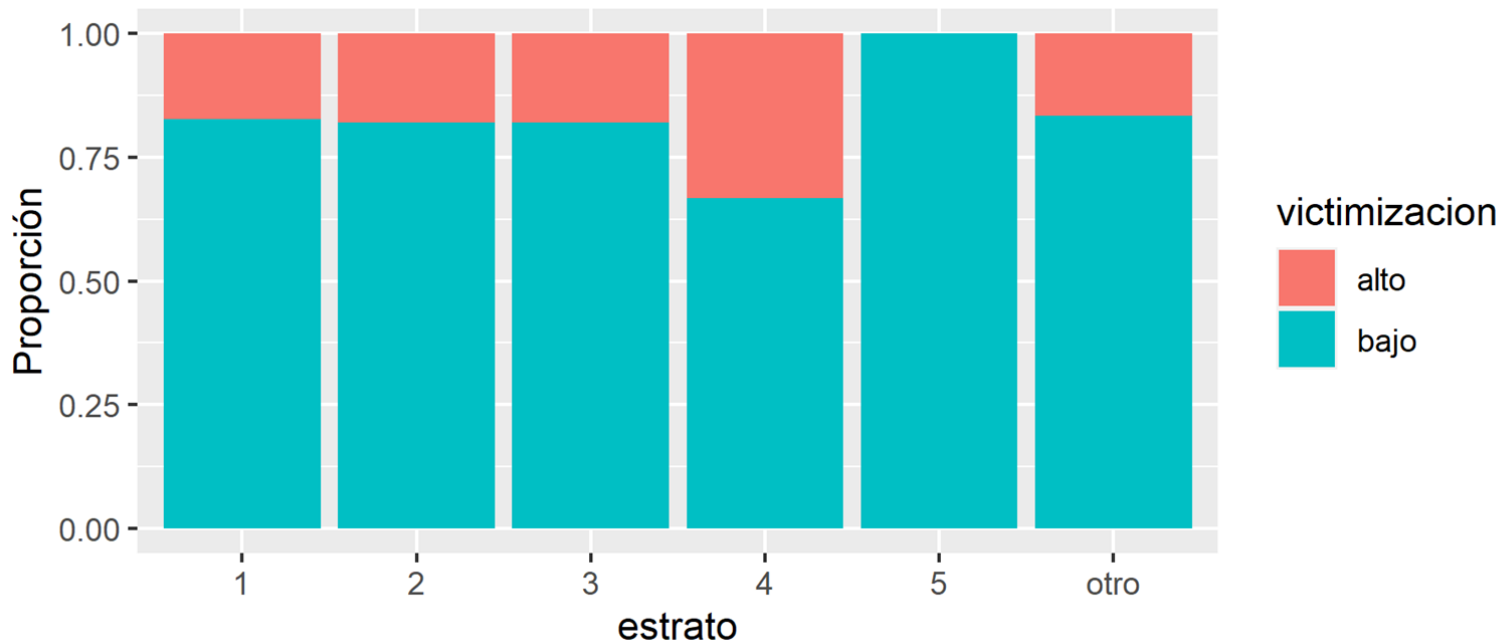
Niveles de violencia por sexo



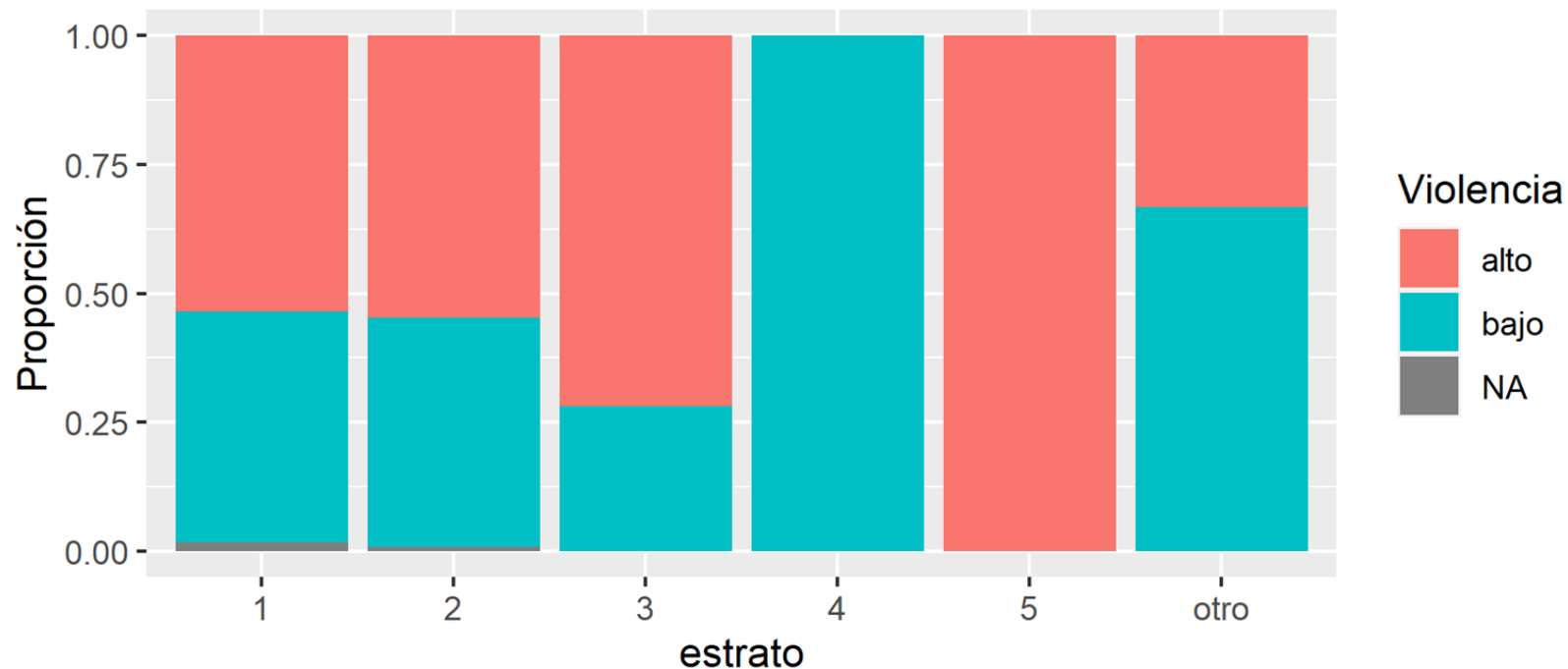
Niveles de ansiedad por estrato



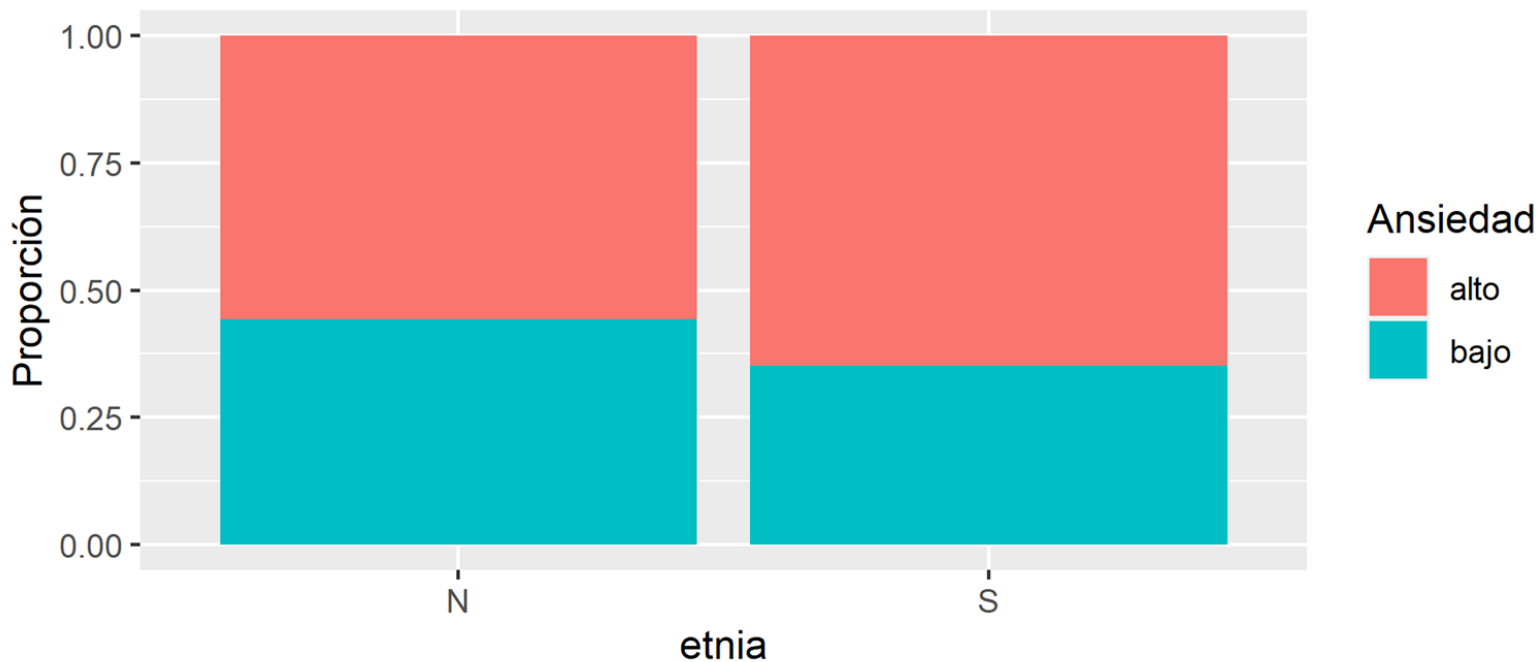
Niveles de victimización por estrato



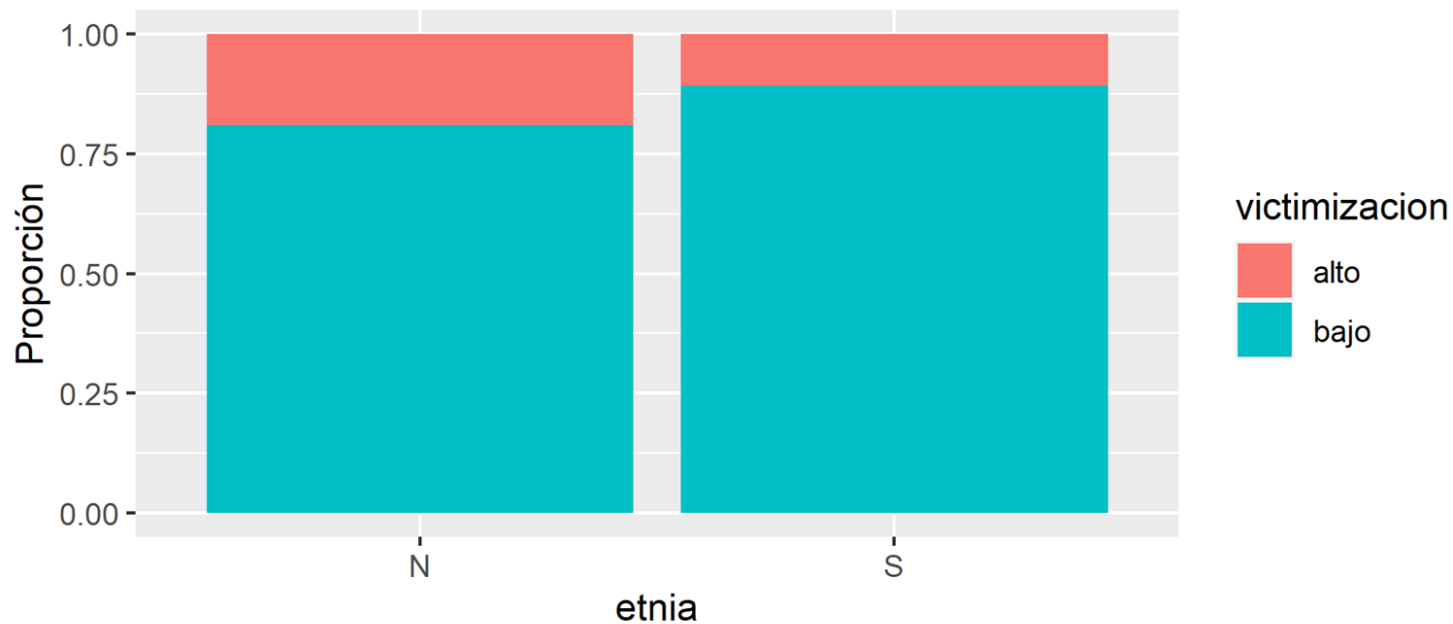
Niveles de violencia por estrato



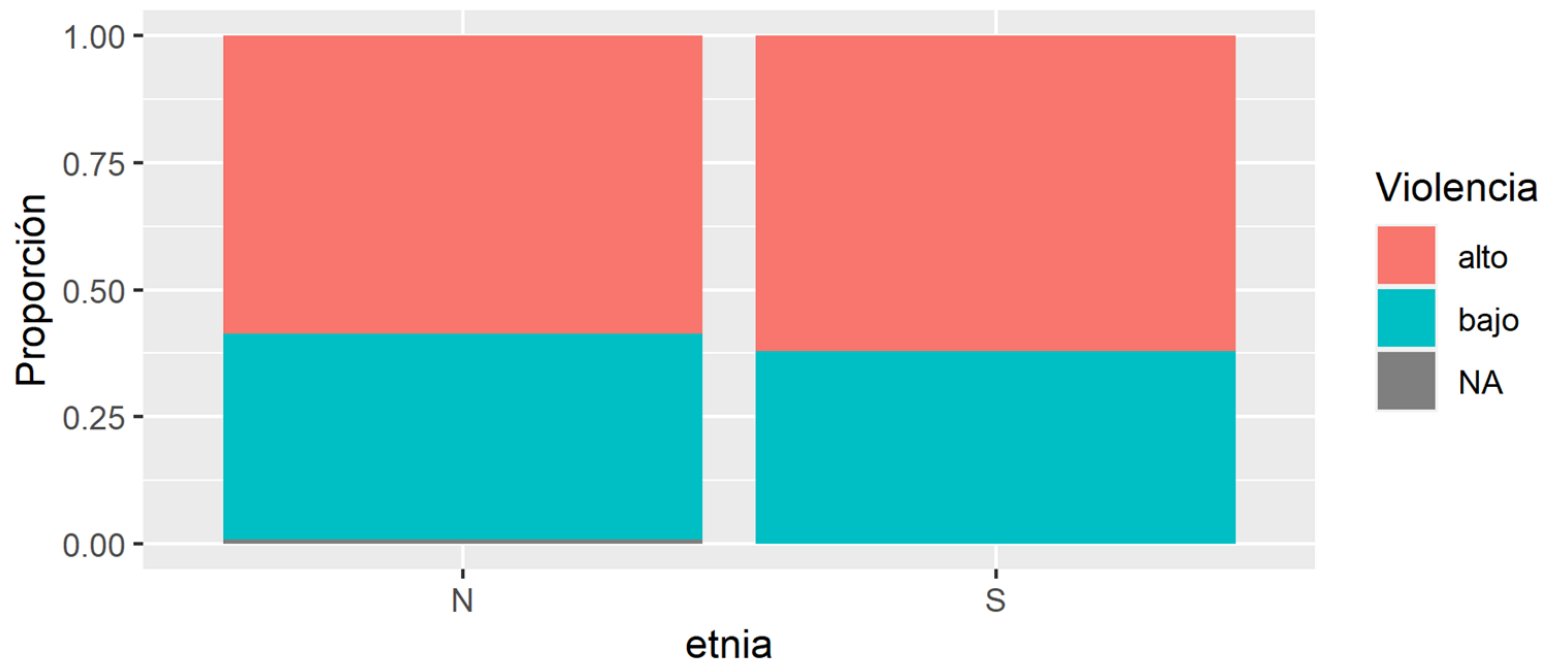
Niveles de ansiedad por etnia



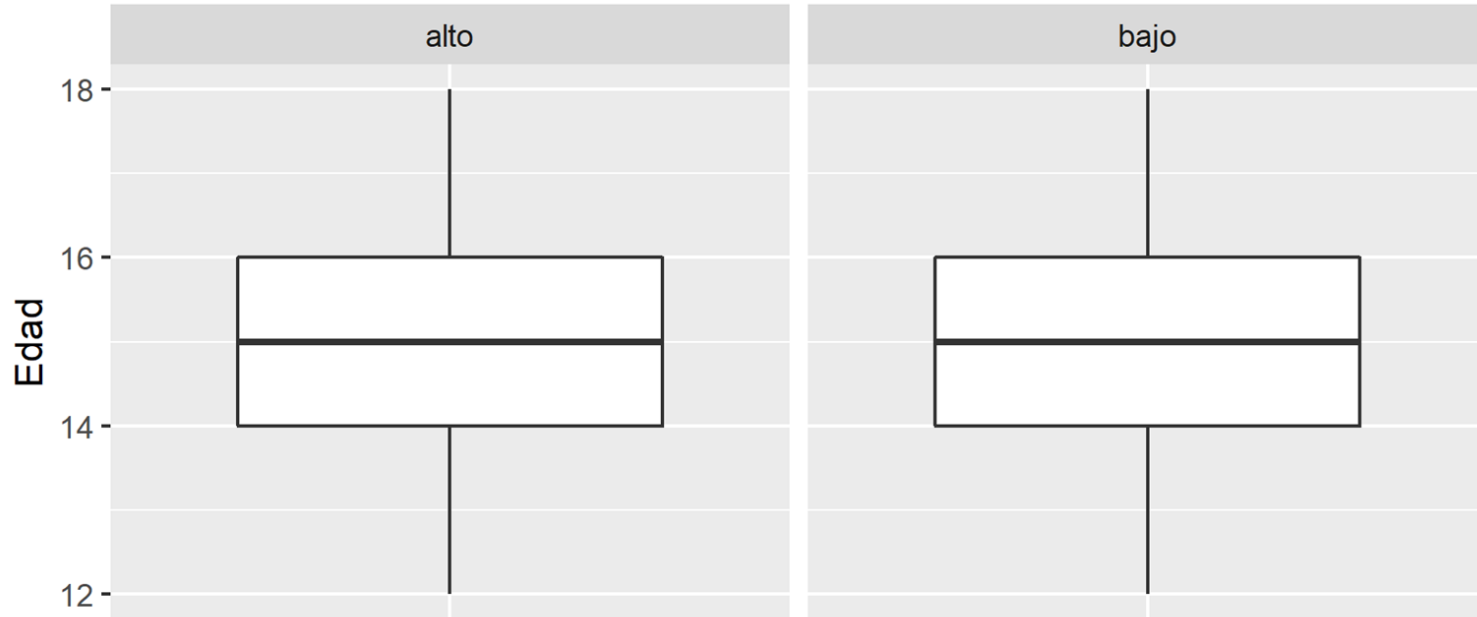
Niveles de victimización por etnia



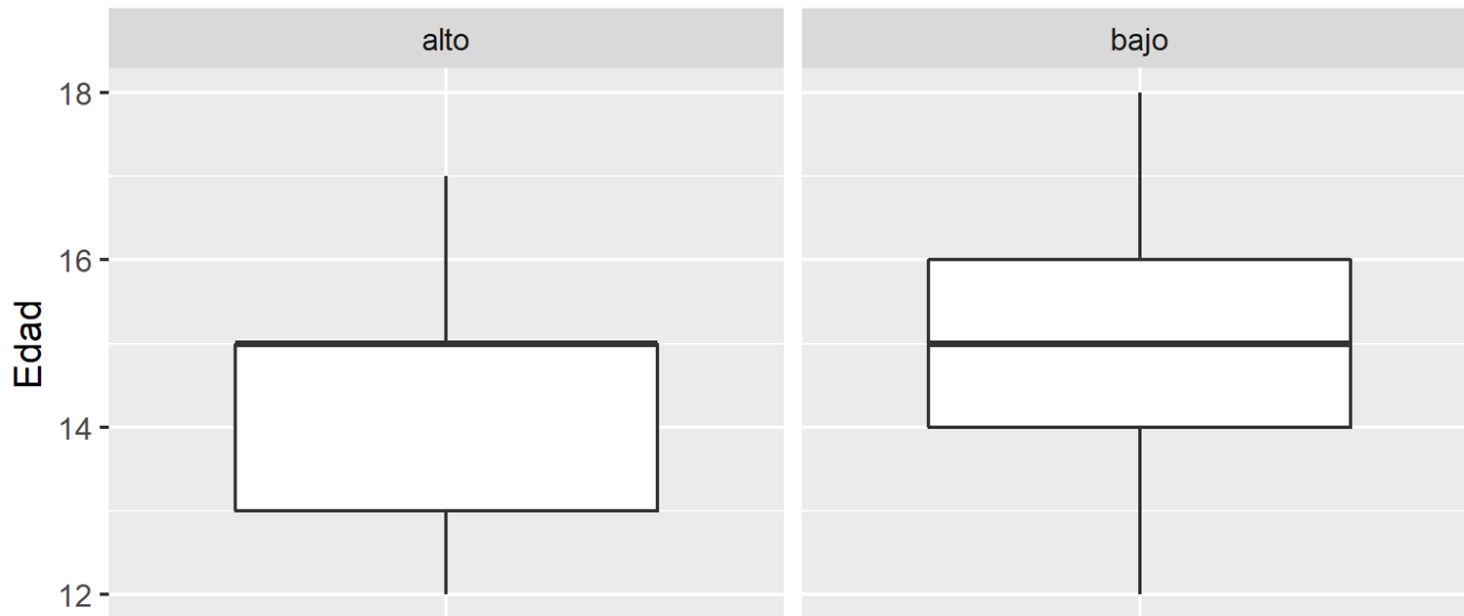
Niveles de violencia por etnia



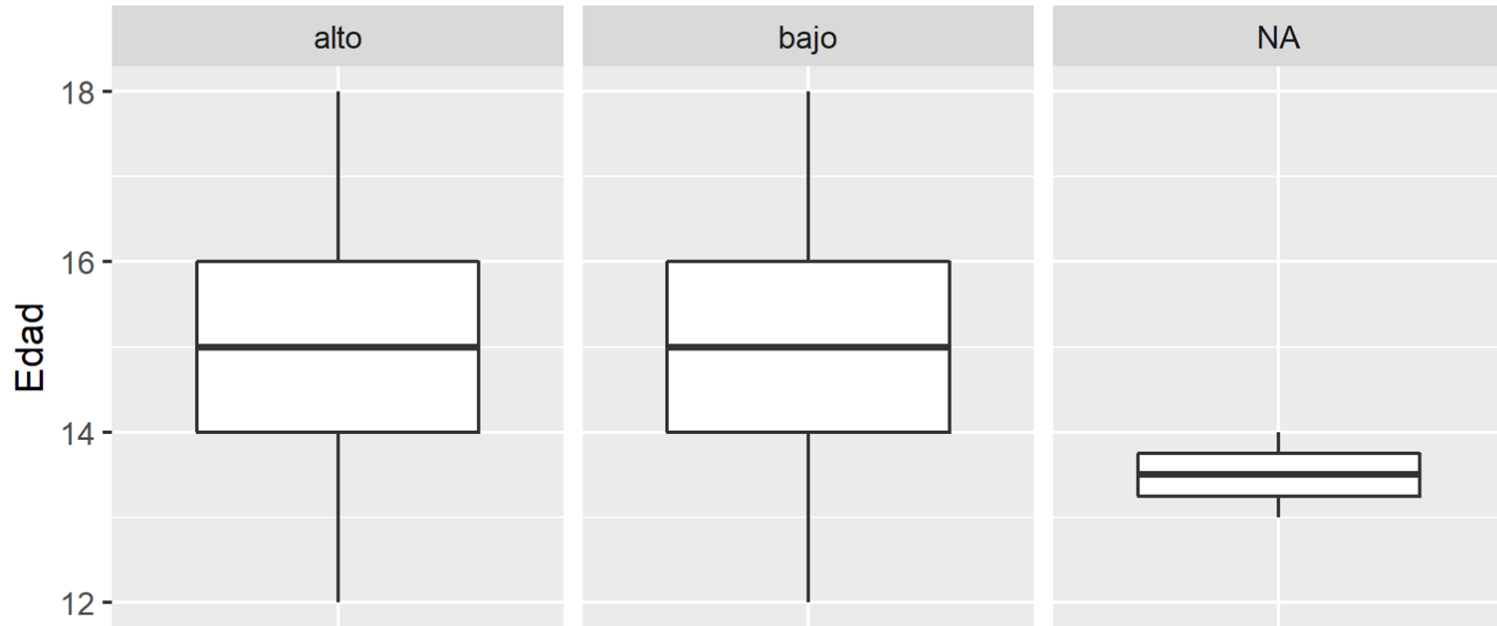
Nivel ansiedad y edad



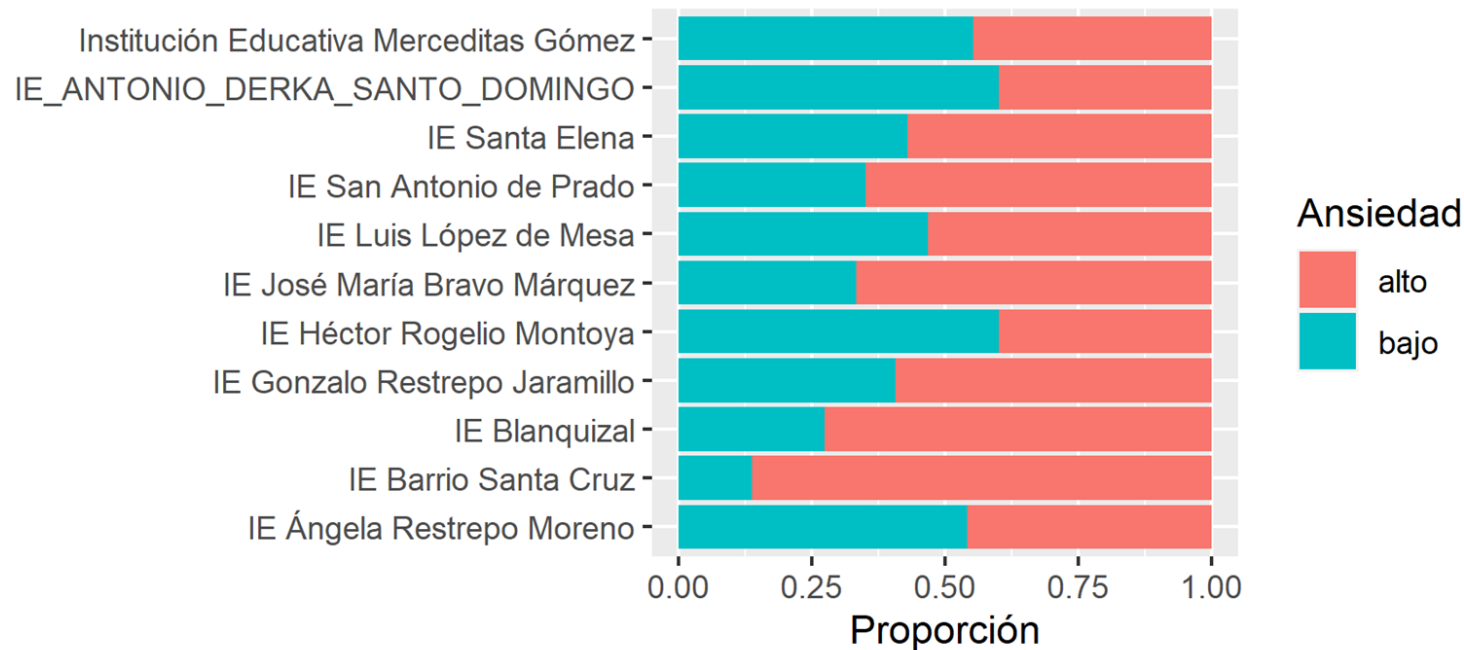
Nivel victimización y edad



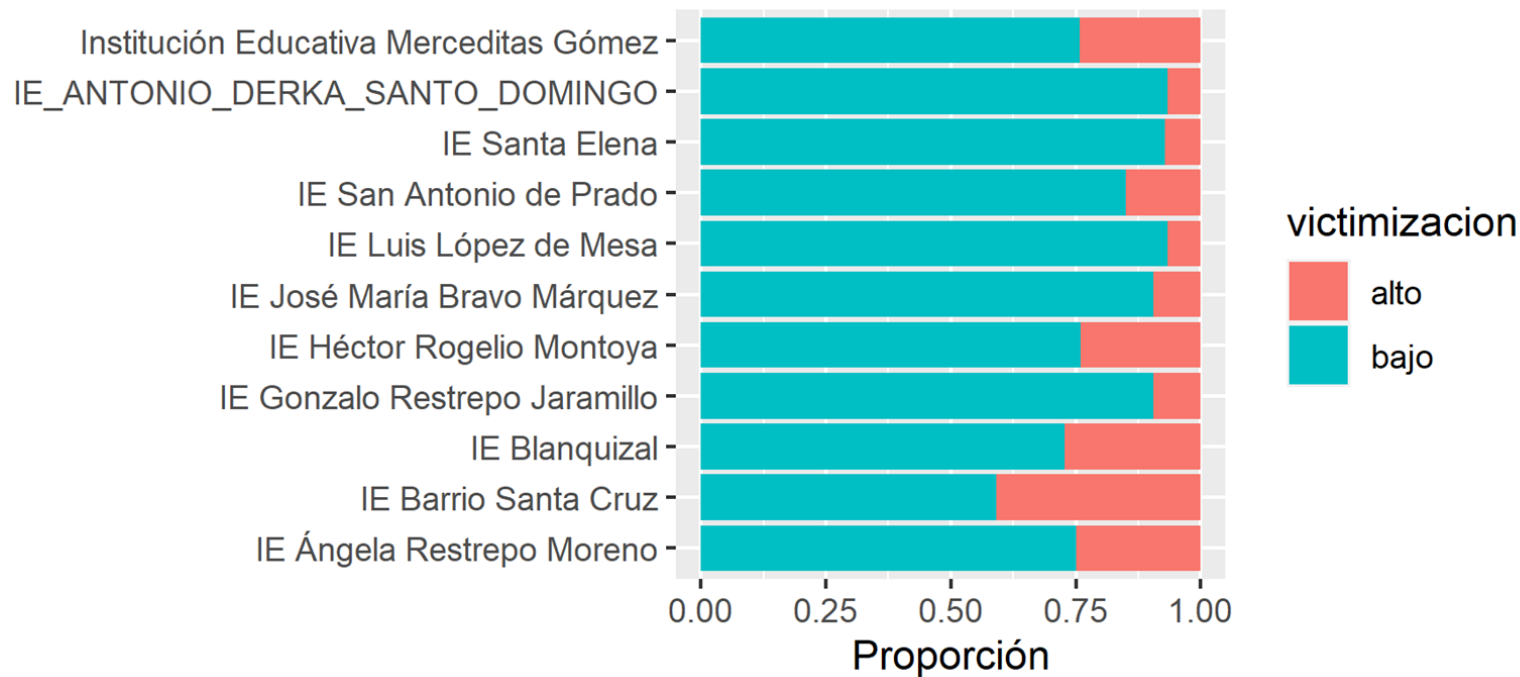
Nivel violencia y edad



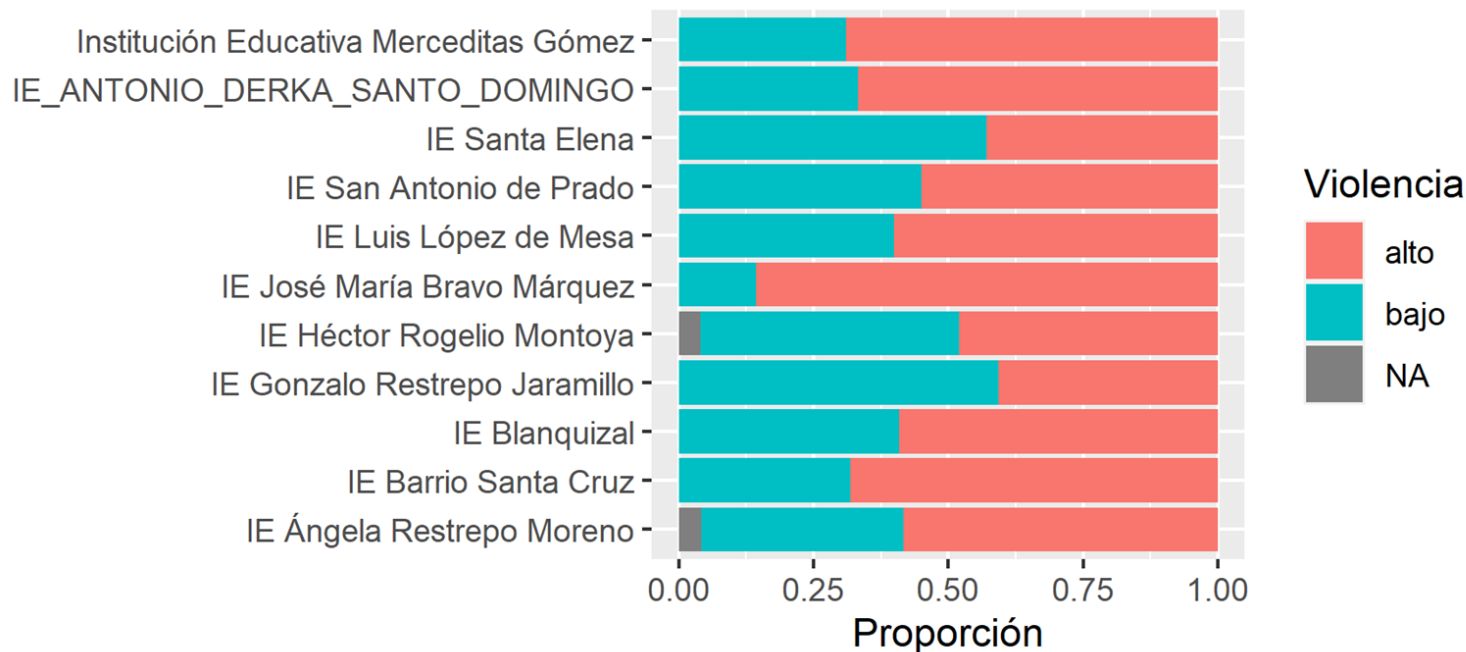
Niveles de ansiedad por institución



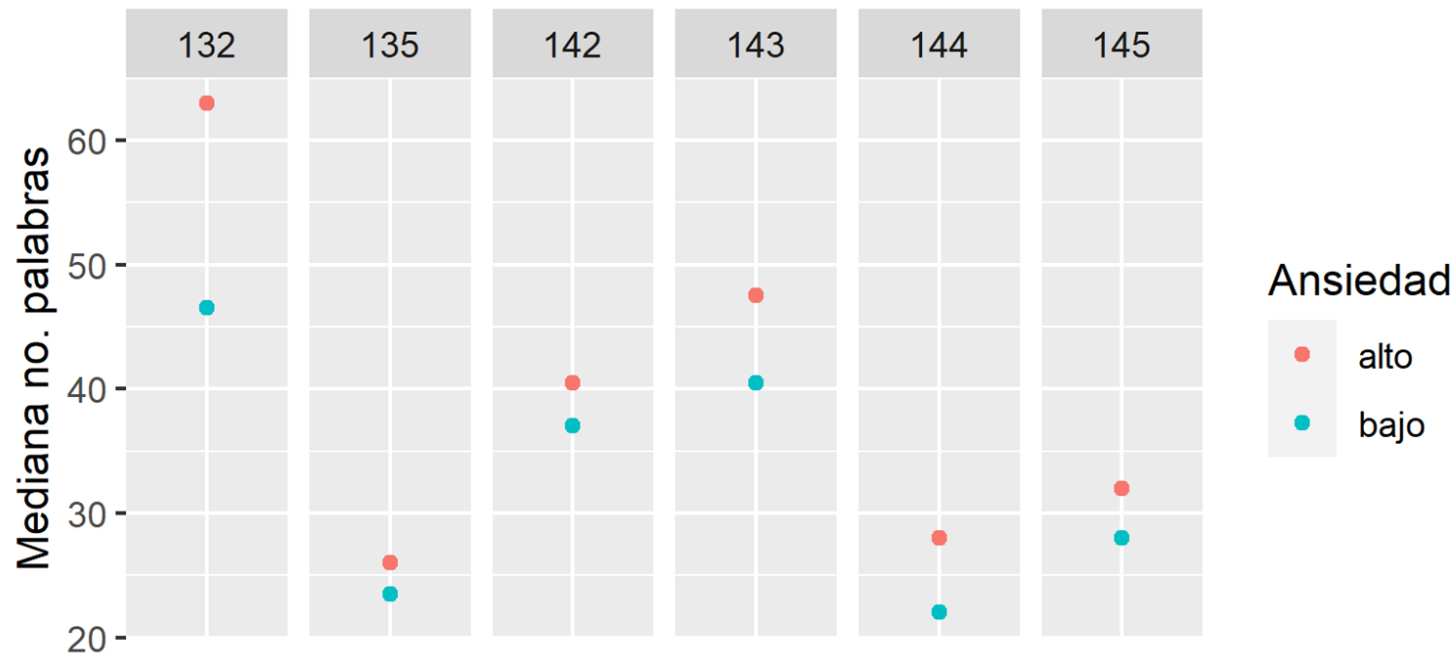
Niveles de victimización por institución



Niveles de violencia por institución



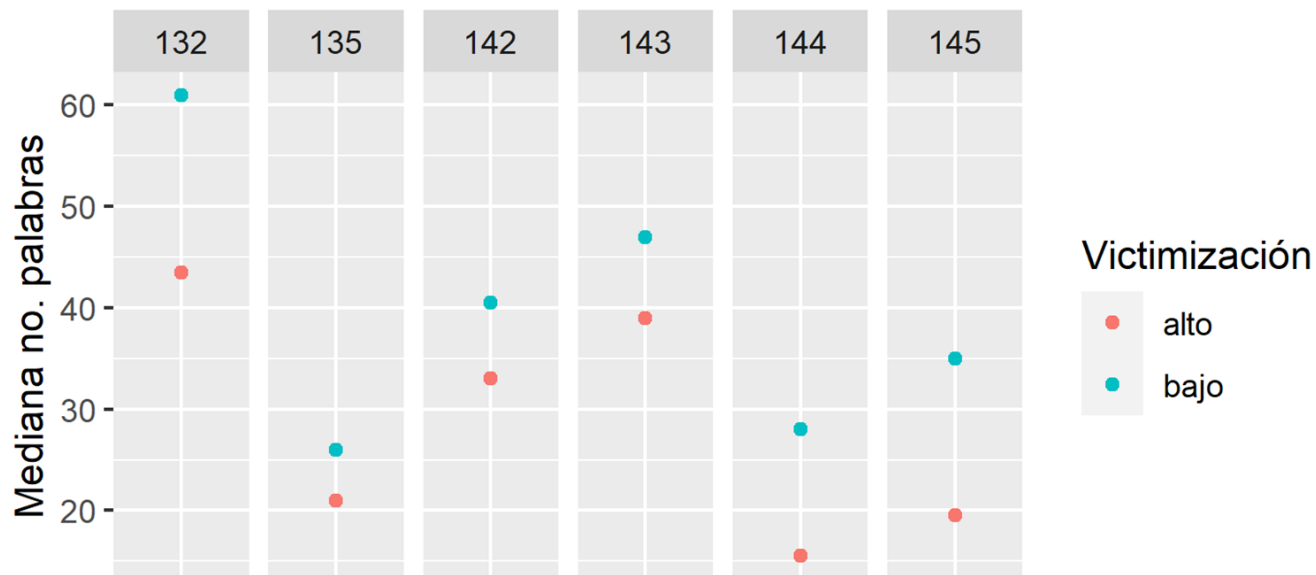
Número de palabras



Número de palabras



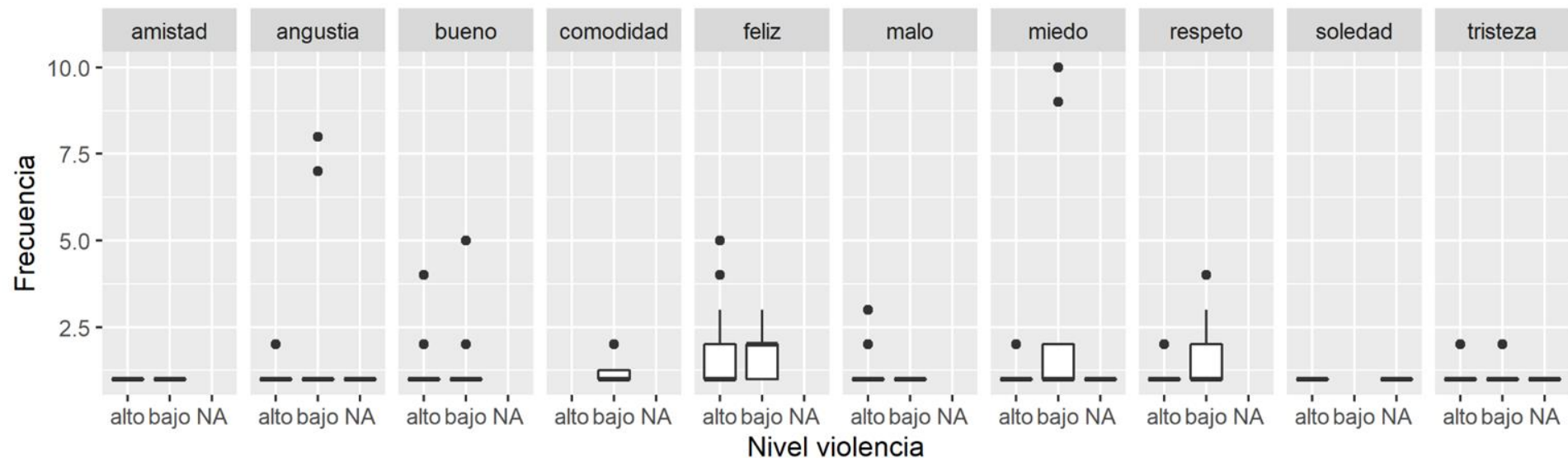
Número de palabras



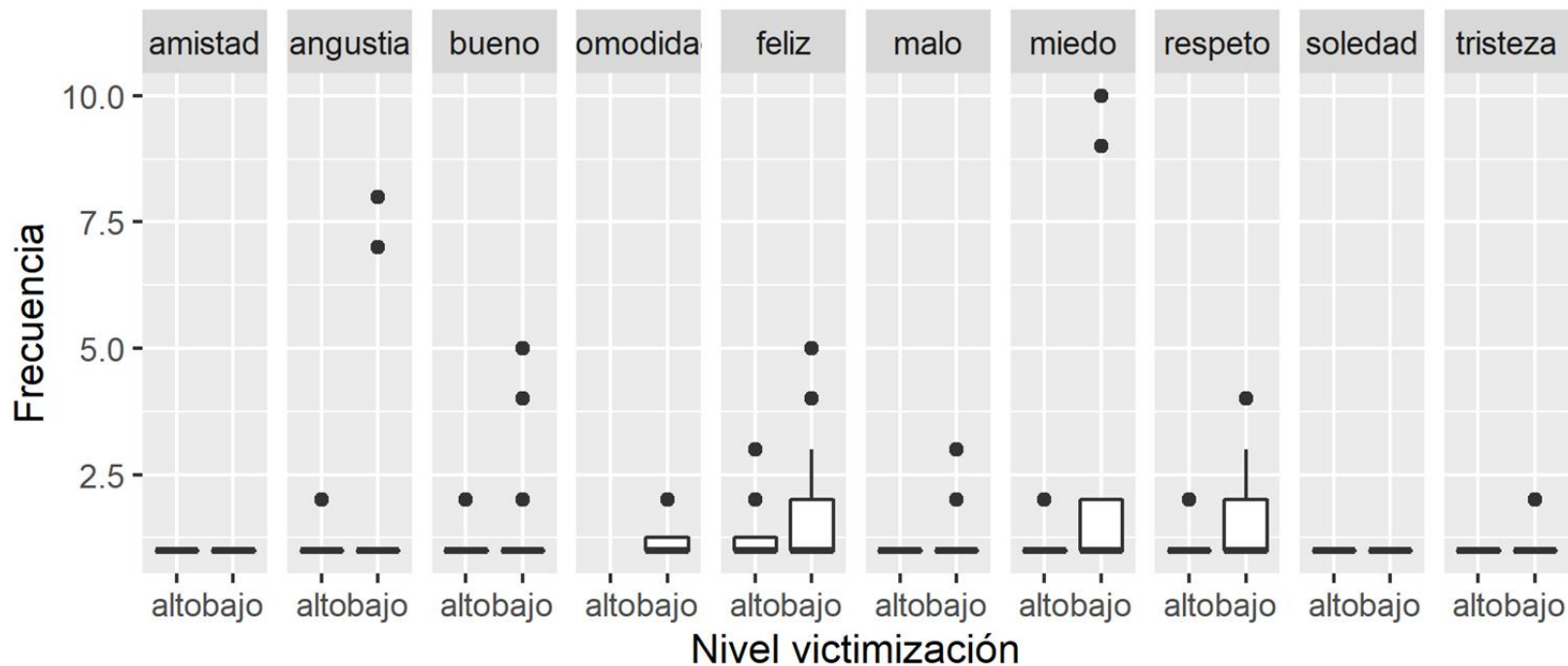
Número de caracteres



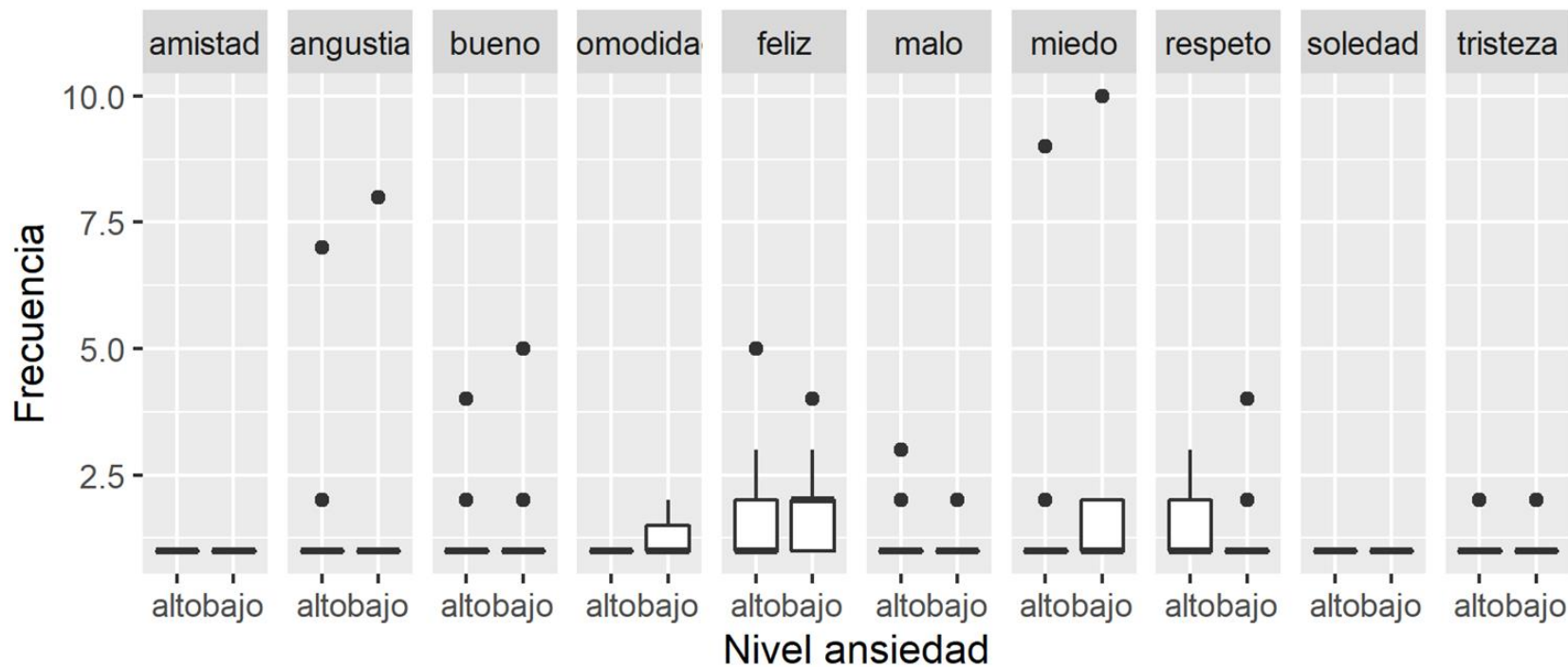
Frecuencia palabras claves



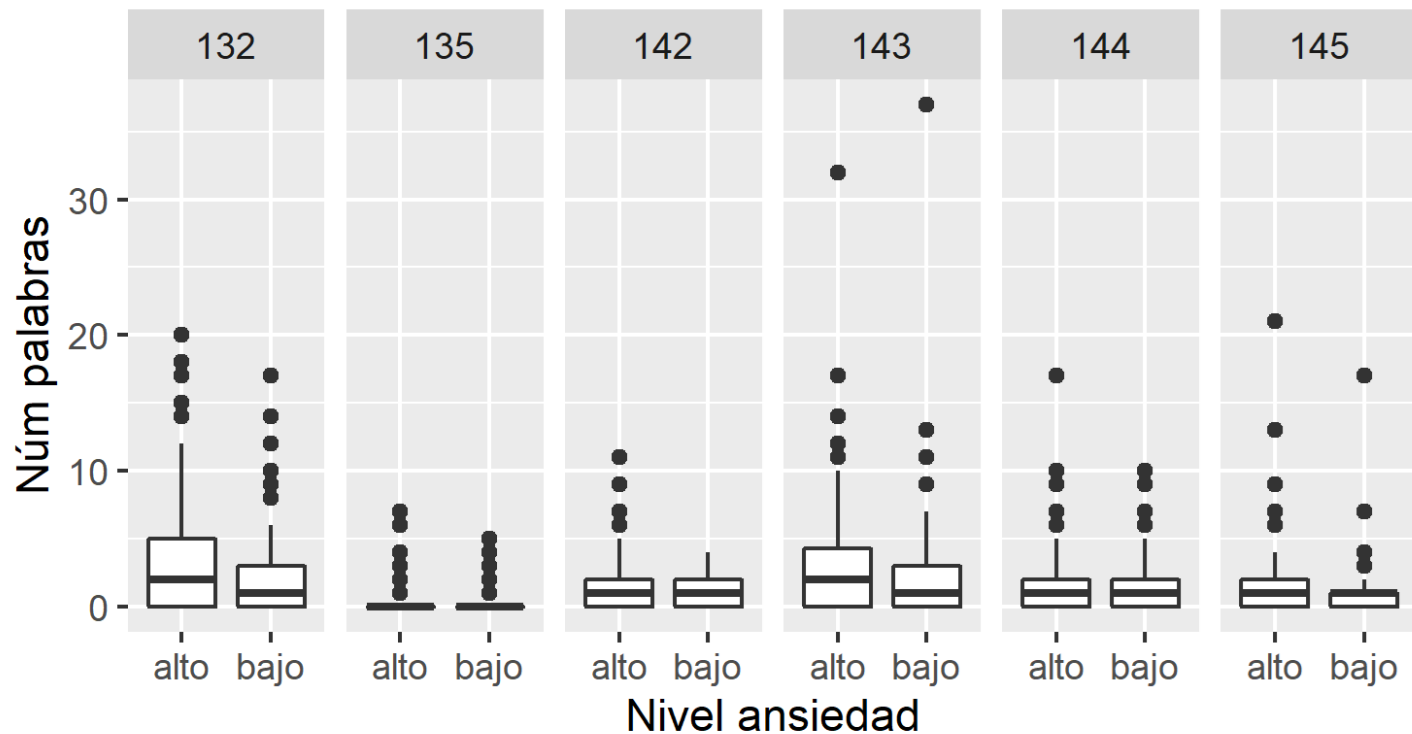
Frecuencia palabras claves



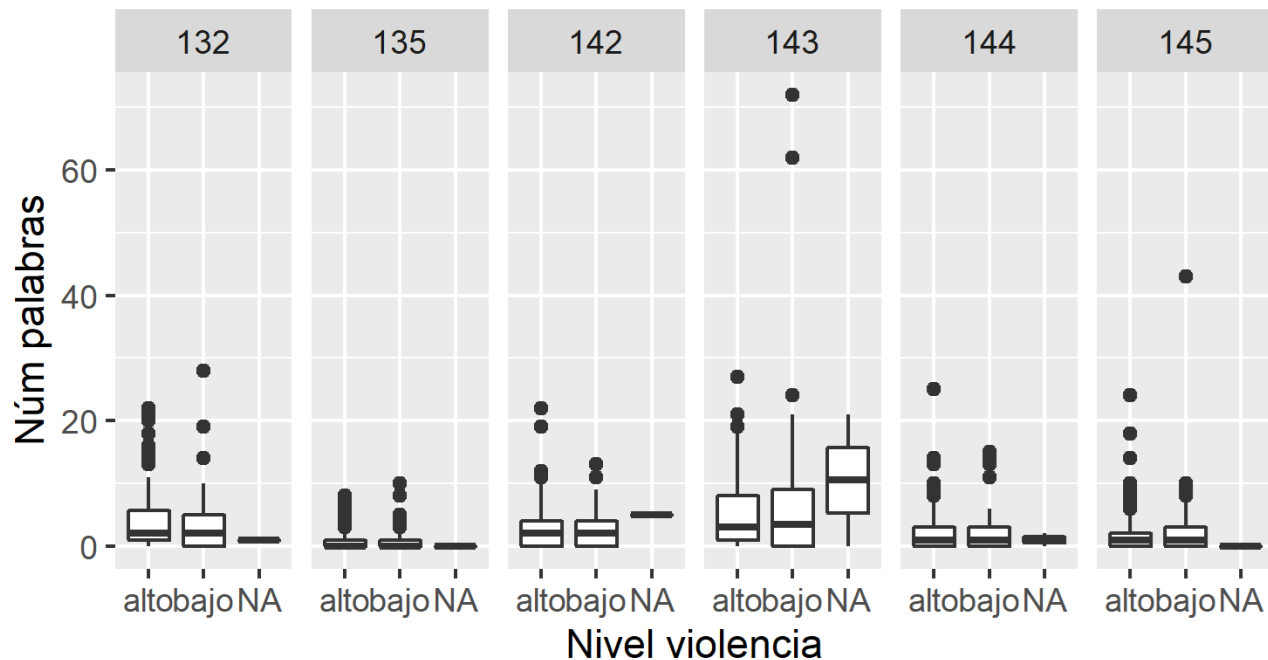
Frecuencia palabras claves



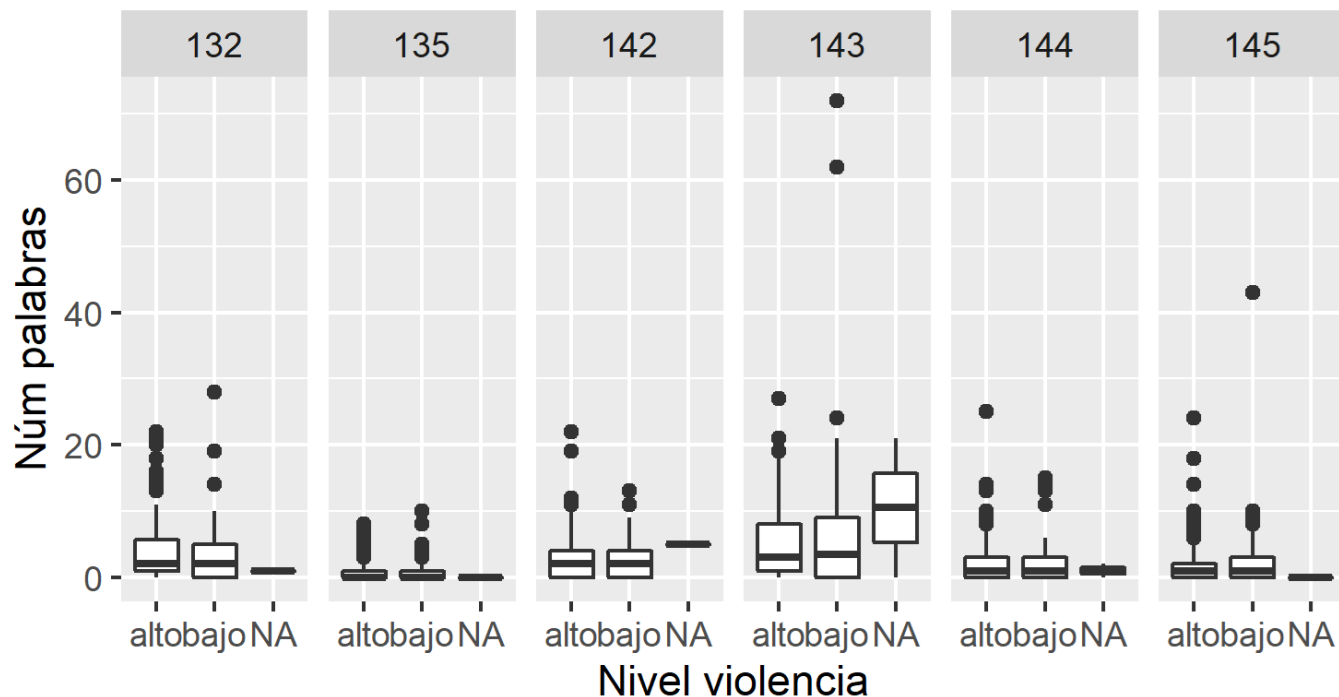
Frecuencia palabras relacionadas con emoción de tristeza



Frecuencia palabras relacionadas con emociones negativas



Frecuencia palabras relacionadas con emoción de tristeza



Frecuencia palabras alerta

