

Diseño y validación de modelos de analítica predictiva para el fenómeno de Riñas y Lesiones Personales y fenómeno de Percepción de Seguridad

Convenio 938 de 2019

Facultad de Ciencias
Universidad Nacional de Colombia

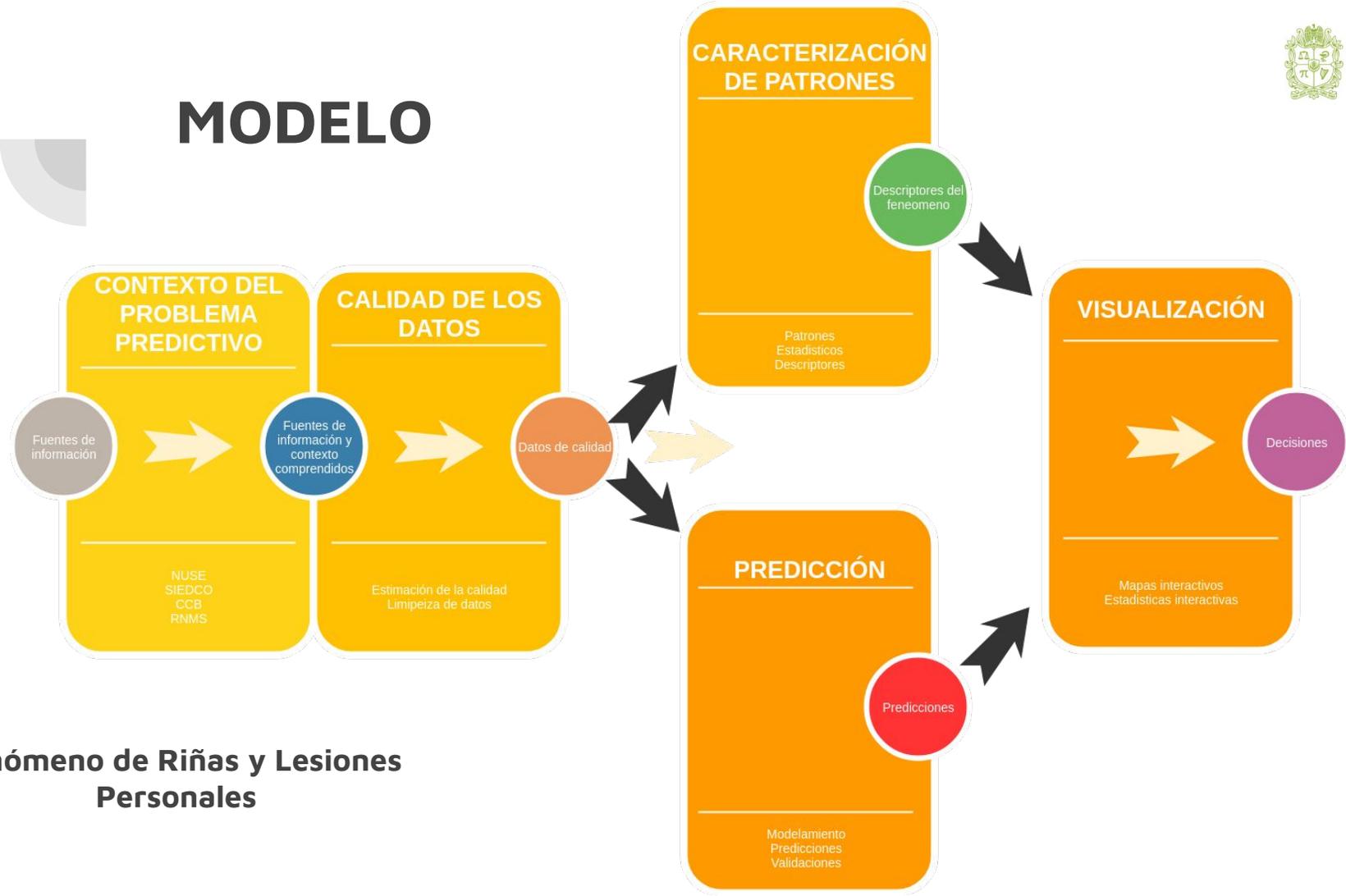
Mayo de 2020



Modelación Fenómeno de Riñas y Lesiones Personales



MODELO



Fenómeno de Riñas y Lesiones Personales



Fenómeno - RIÑAS



Aproximación

- Caracterización de asociaciones secundarias entre la ocurrencia de las riñas y factores socioeconómicos
- Patrones temporales de la ocurrencia de riñas
- Modelo predictivo espacio temporal de la ocurrencia de las riñas

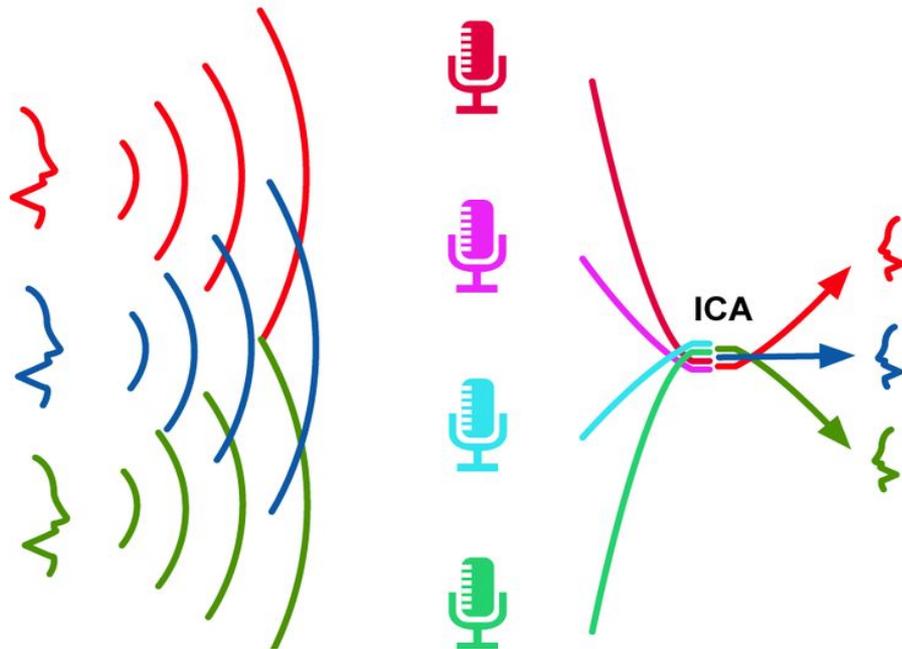




Descomposición espacial de la ocurrencia de riñas

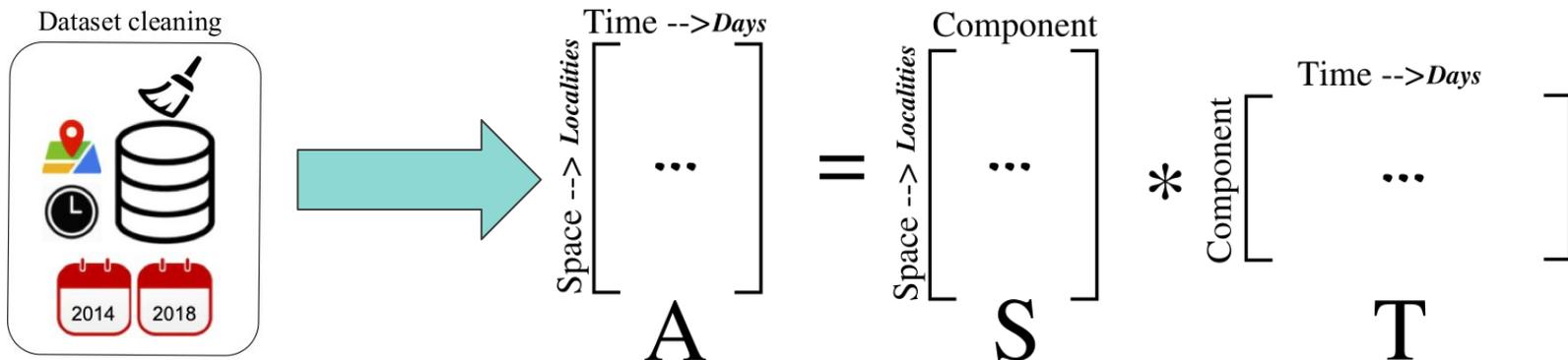
Fenómeno de Riñas y Lesiones Personales (Aproximación # 1)

Pregunta de investigación: ¿Existen factores latentes que subyacen a la ocurrencia de las riñas en la ciudad de Bogotá?



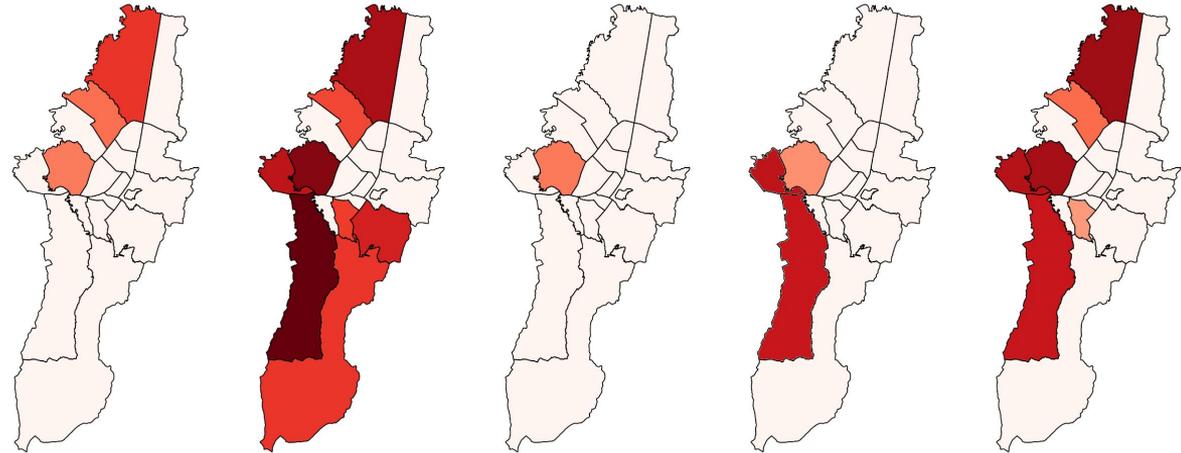
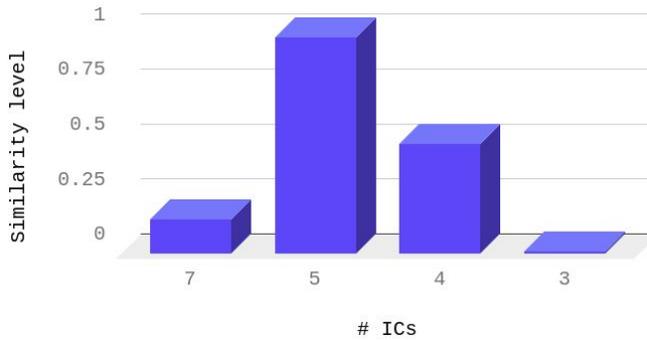
Fenómeno de Riñas y Lesiones Personales (Aproximación # 1)

Pregunta de investigación: ¿Existen factores latentes que subyacen a la ocurrencia de las riñas en la ciudad de Bogotá?

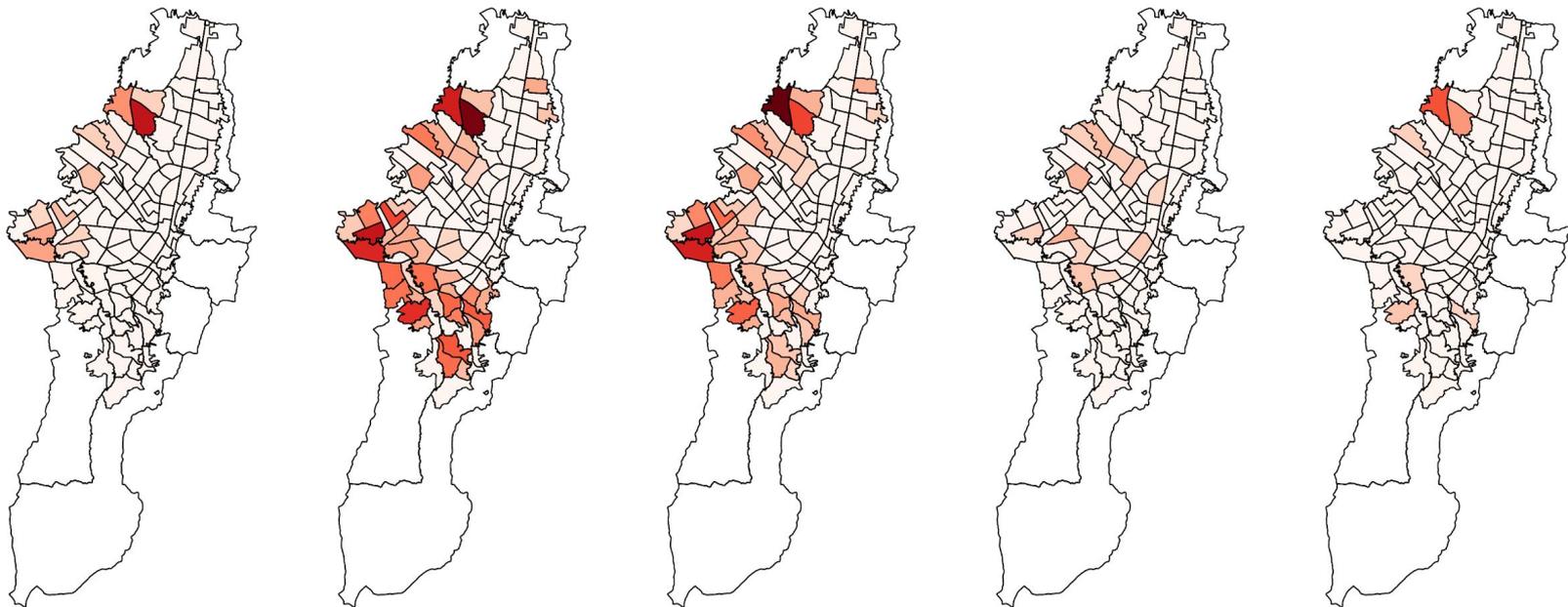


Análisis de componentes independientes

Resultados - Aproximación # 1



Resultados - Aproximación # 1



Mensajes:

- 5 factores parecen subyacer a la ocurrencia de las riñas en la ciudad de Bogotá. Un análisis más profundo es requerido para comprender la naturaleza de esto.
- Usme, Ciudad Bolívar y San Cristóbal parecen compartir un dinámica temporal de ocurrencia de riñas común. Suba y Engativá también parecen compartir una dinámica común.

Trabajo Futuro:

- Explorar la relación entre los componentes y variables socio-económicas.

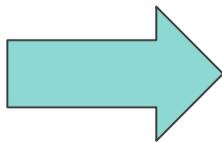


Asociaciones secundarias entre la ocurrencia de riñas y variables socioeconómicas



Fenómeno de Riñas y Lesiones Personales (Aproximación # 2)

Pregunta de investigación: ¿Existen asociaciones secundarias entre la ocurrencia de las riñas y diversas variables socioeconómicas en la ciudad de Bogotá?



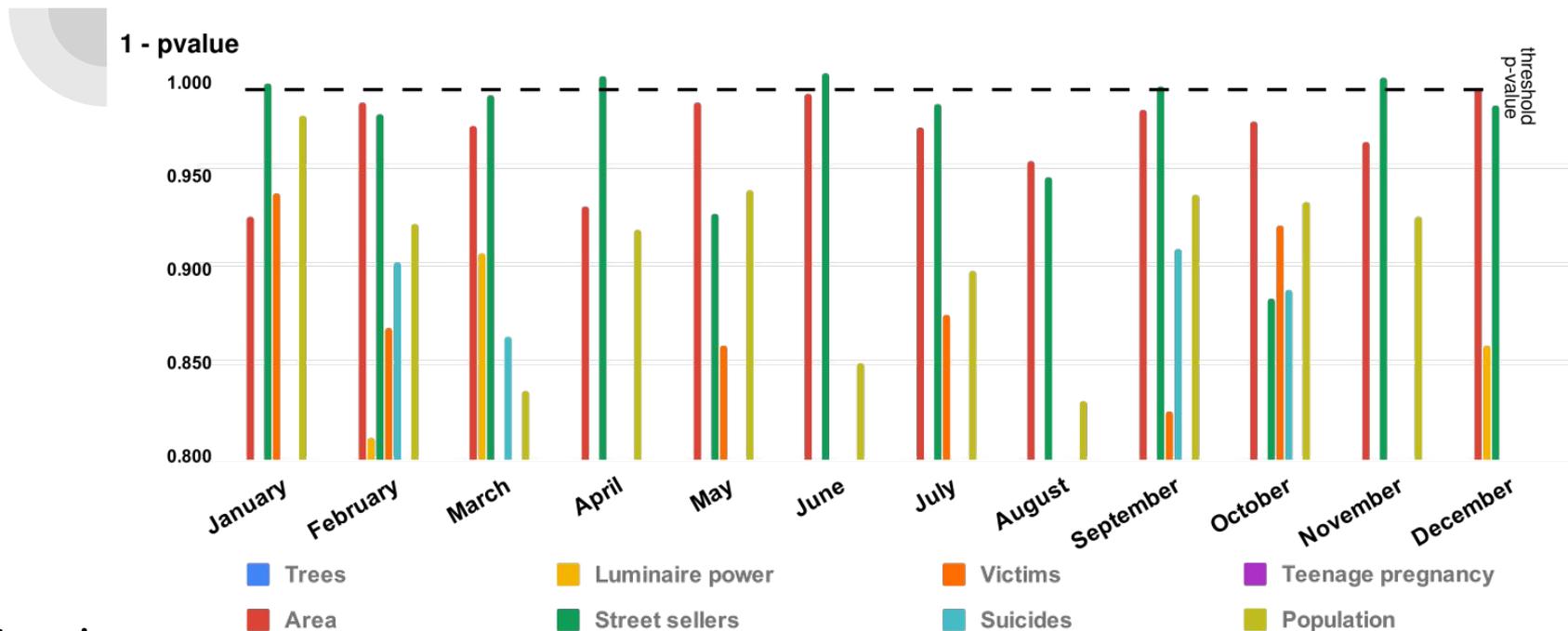
$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i$$
$$Y = \beta_0 + \beta X + \varepsilon$$

General Linear model (GLM)

Resultados - Aproximación # 2

Factor	Coef	Std error	z	P> z
Number of trees	0.0804	0.176	0.457	0.648
Area	-0.4213	0.181	-2.325	0.02 *
Luminaire power of public lighting	-0.4361	0.384	-1.136	0.256
Number of street sellers	0.6333	0.253	2.502	0.012 *
Number of victims of the armed conflict	0.244	0.197	1.236	0.217
Number of suicides	0.3777	0.352	1.073	0.283
Number of teenage pregnancy	0.1178	0.343	0.344	0.731

Resultados - Aproximación # 2



Mensajes:

- El número de vendedores ambulantes pare explicar significativamente la ocurrencia de riñas en la ciudad de Bogotá durante el año 2018.

Trabajo Futuro:

- Explorar otras variables socio-económicas con potencial explicativo y una escala espacial diferente.



Cuantificación de los niveles de predictibilidad de las riñas



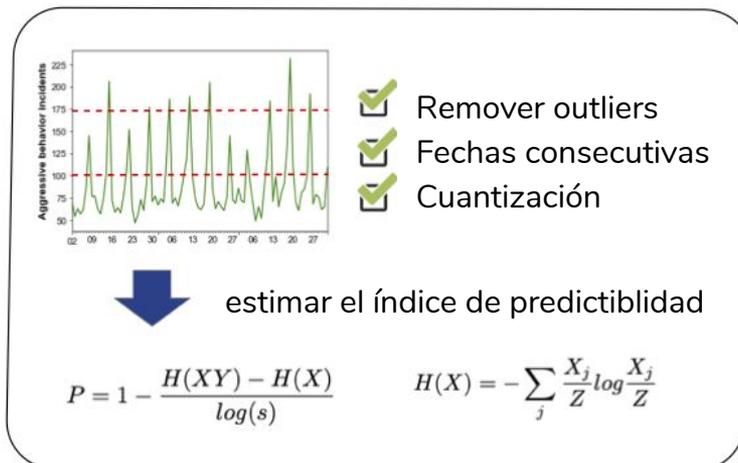
Fenómeno de Riñas y Lesiones Personales (Aproximación # 3)

Pregunta de investigación: ¿La ocurrencia de comportamientos agresivos (riñas) en la ciudad de Bogotá exhibe patrones temporales que pueden emplearse para identificar el nivel de predictibilidad de riñas?

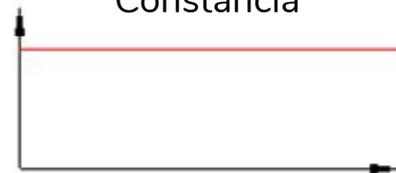
Limpieza dataset



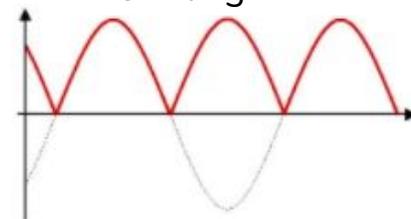
Pre-procesamiento + medición predictibilidad



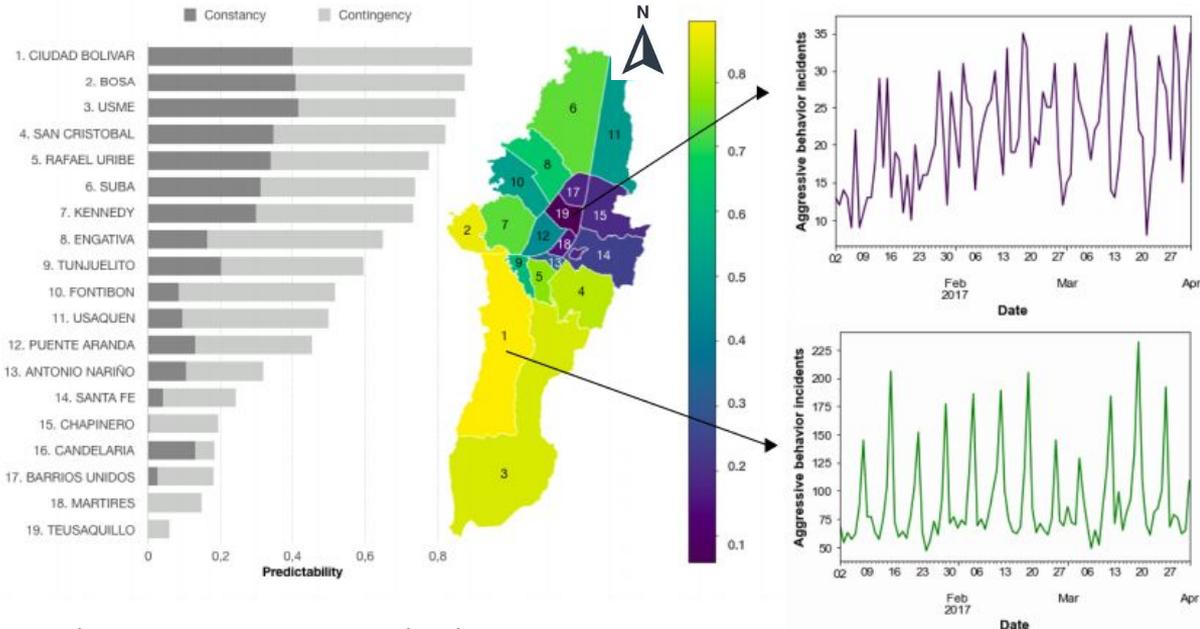
Constancia



Contingencia



Resultados - Aproximación # 3



- Es posible **caracterizar** de manera **cuantitativa** el nivel en que la ocurrencia de riñas exhibe **patrones temporales** en la ciudad de Bogotá a través del **índice de predictibilidad**.
- El nivel de **predictibilidad no es homogéneo** en la ciudad.
- Las **localidades del sur** de la ciudad son **más predecibles**.
- Las localidades del centro y el oriente, son menos predecibles.
- Se observa una **mayor contribución de la contingencia** sobre el índice de predictibilidad para la mayoría de localidades.

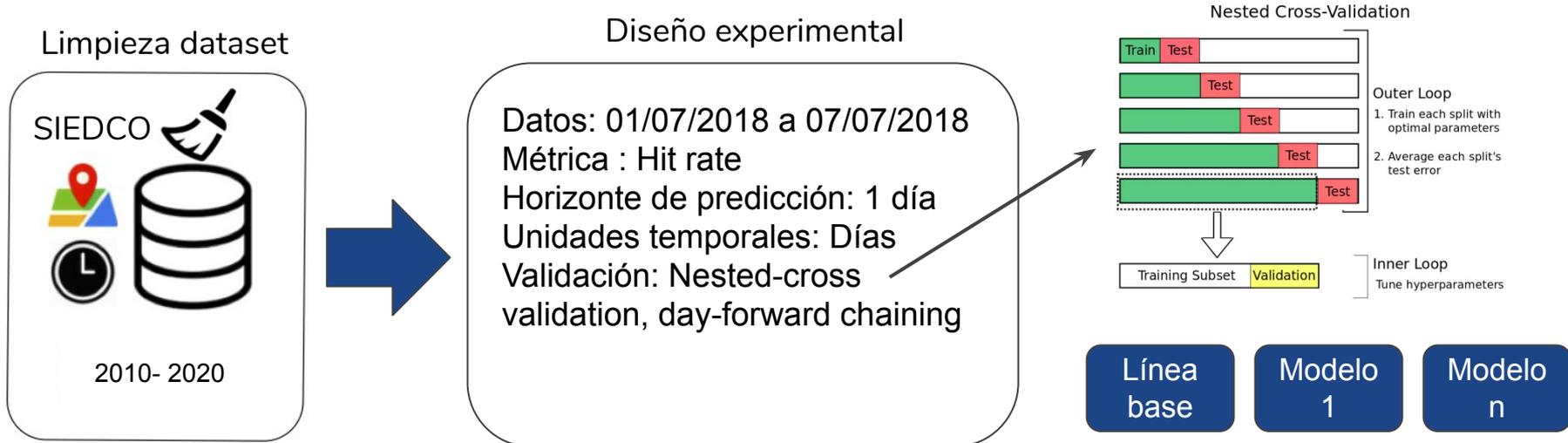


Predicción espacio-temporal de la ocurrencia de riñas



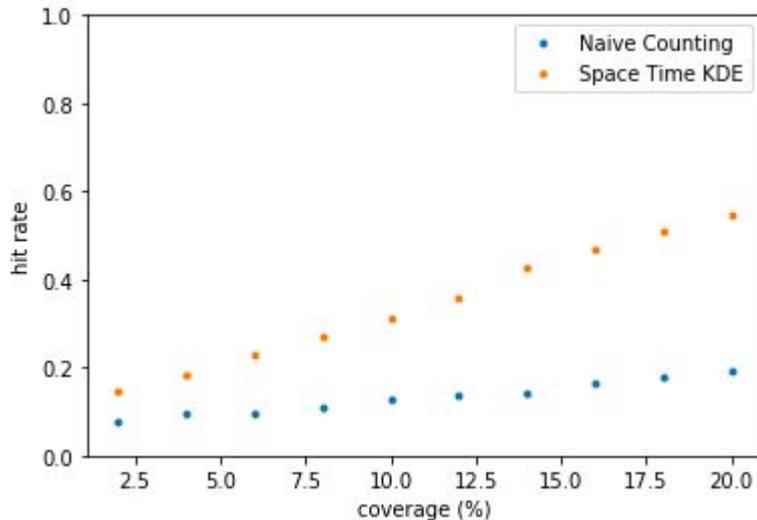
Fenómeno de Riñas y Lesiones Personales (Aproximación # 4)

Pregunta de investigación: ¿Es posible predecir (en el corto plazo, i.e. día) la ocurrencia espacio temporal de riñas en la ciudad de Bogotá?



Resultados - Aproximación # 4

Pregunta de investigación: ¿Es posible predecir (en el corto plazo, i.e. día) la ocurrencia espacio temporal de riñas en la ciudad de Bogotá?



Modelo línea base:
Naive Counting



Hit rate
(cobertura 20%)

0.190610

Modelo 2: Space
Time KDE

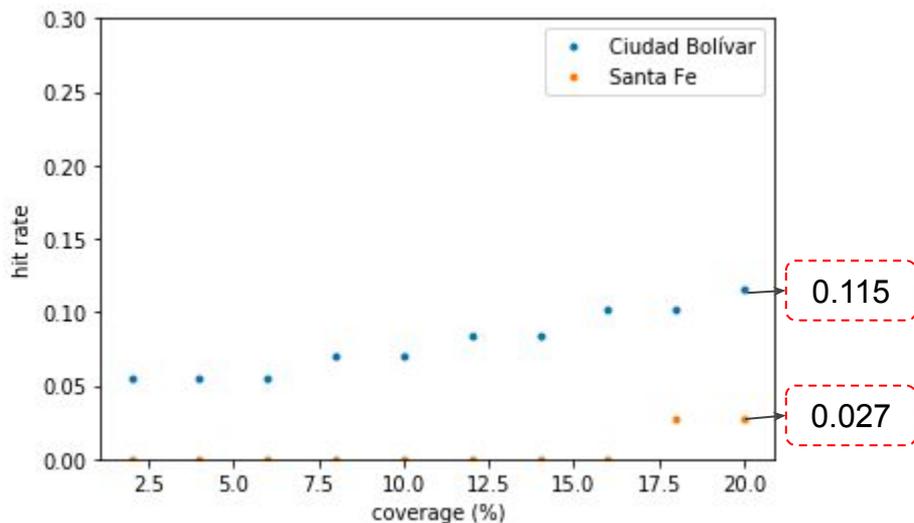


0.544586

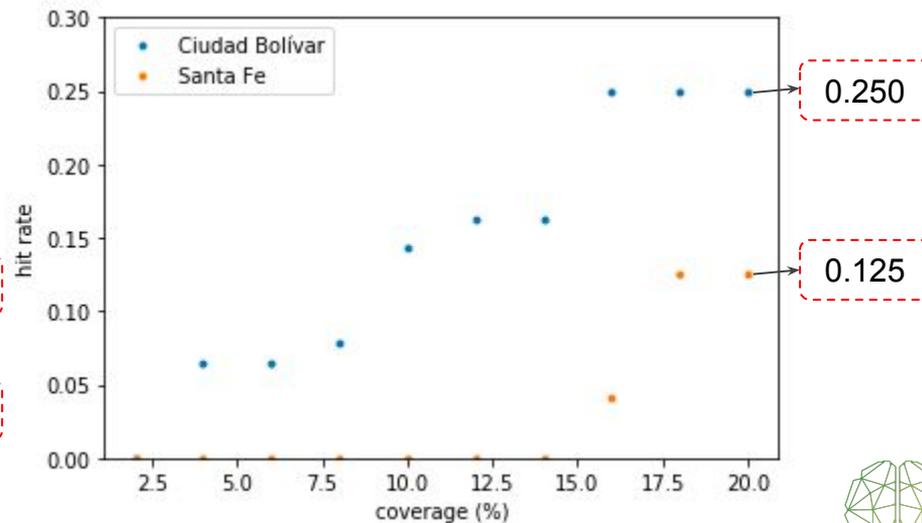
Resultados - Aproximación # 4

Pregunta de investigación: ¿Es posible predecir (en el corto plazo, i.e. día) la ocurrencia espacio temporal de riñas en la ciudad de Bogotá?

Modelo línea base: Naive Counting



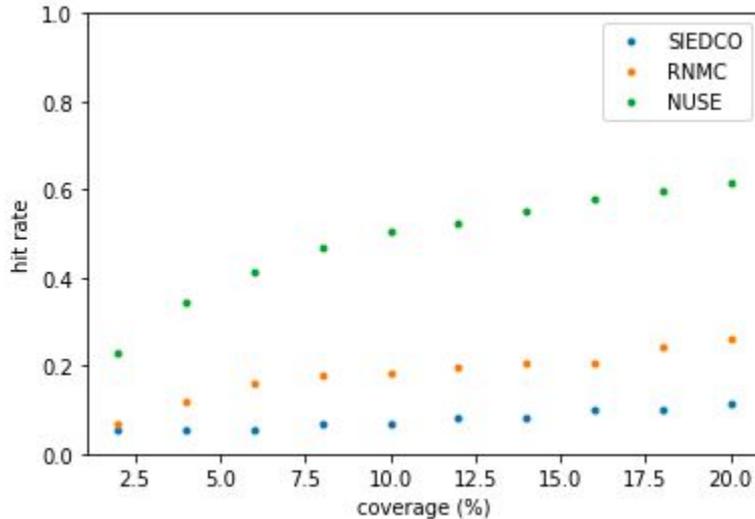
Modelo 2: Space Time KDE



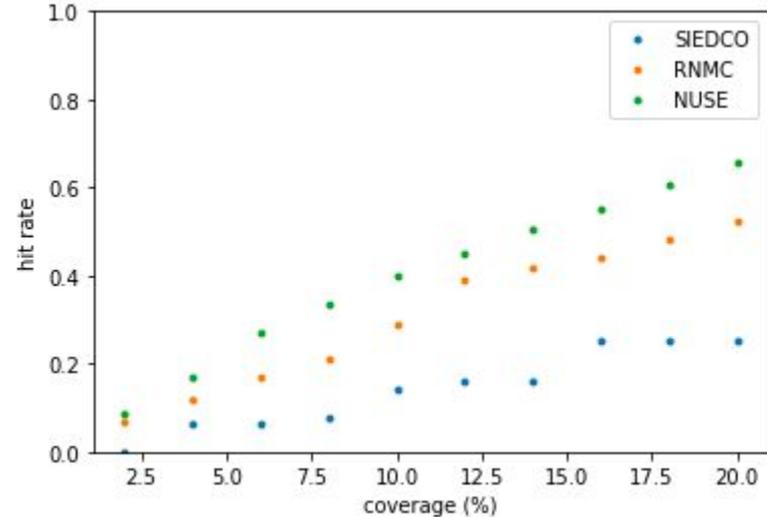
Resultados - Aproximación # 4

Pregunta de investigación: ¿Es posible predecir (en el corto plazo, i.e. día) la **ocurrencia** espacio temporal de riñas en la ciudad de Bogotá?

Modelo línea base: Naive Counting



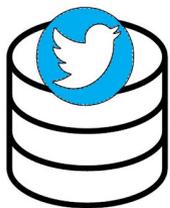
Modelo 2: Space Time KDE





Fenómeno - Percepción de la seguridad

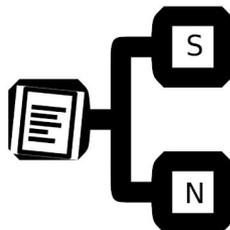
Modelo propuesto



Recolección
datos de
Twitter



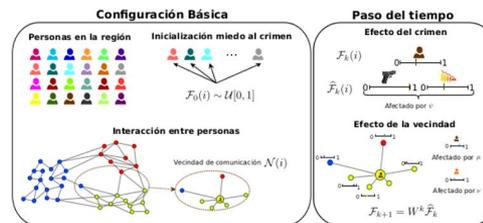
Filtrado
textos
relacionados
Seguridad



Clasificación
textos que
hablan
de seguridad

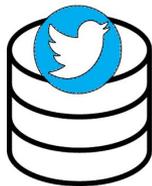


Medición
percepción
seguridad



Aplicación modelo
de propagación
de opinión seguridad

Fenómeno de Percepción de Seguridad



Recolección
datos de
Twitter

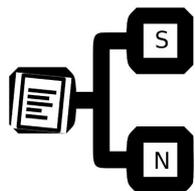
Periodo:
18/12/2019 -- 20/04/2020
~2000 tweets/hora
4 millones de tweets
75% son retweets



Filtrado
textos
relacionados
Seguridad

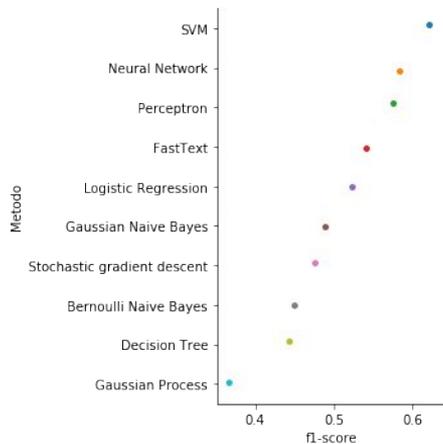
Listado de palabras clave
compuestas proporcionado
por la SSCJ.

~7500 Tweets únicos



Clasificación
textos que
hablan
de seguridad

Clasificación manual.
Ajuste automático con
modelos de aprendizaje de
máquina.



Matriz de Confusión

True label	Predicted label	
	No Seguridad	Seguridad
No Seguridad	207	33
Seguridad	16	40

Predicted label
f1_score=0.6202





Análisis de sentimientos de percepción de seguridad en redes sociales



Fenómeno de Percepción de Seguridad

Análisis de Sentimientos

Detección de emociones, como la felicidad, la frustración, la ira, la tristeza, etc.

¿Que tan fina se desea la polaridad?
(continua o discreta)

Lexicones (y métodos tradicionales de ML) o Algoritmos de Deep Learning



"La Seguridad
ha mejorado
en mi sector"

POSITIVO



"Fue un regreso
normal a casa,
TM estuvo OK"

NEUTRAL



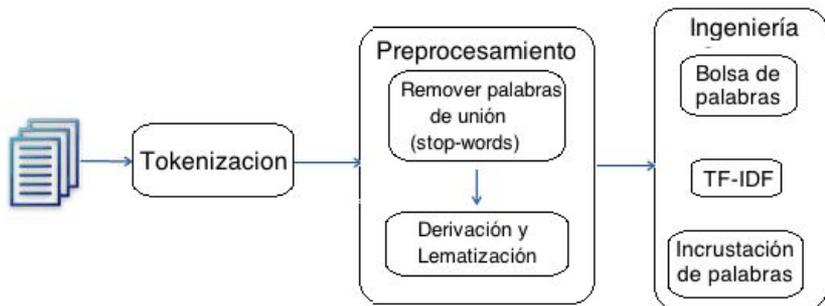
"Que inseguridad,
¿donde está la
policía cuando
se necesita?"

NEGATIVO

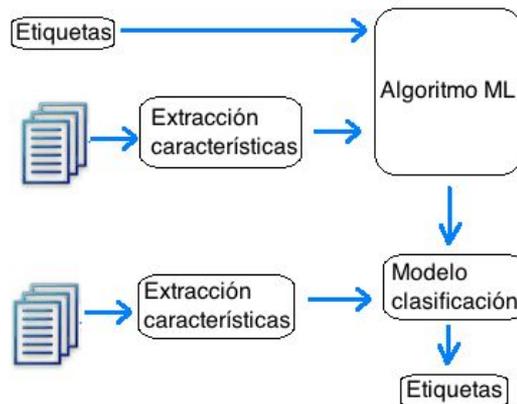
Fenómeno de Percepción de Seguridad

Análisis de Sentimientos

Basados en reglas:
reglas para identificar
subjetividad
conteo de palabras



Aprendizaje automático:
Extracción de características
para entrenamiento de
modelo





Predicción de la propagación de la percepción de la seguridad

Fenómeno de Percepción de Seguridad

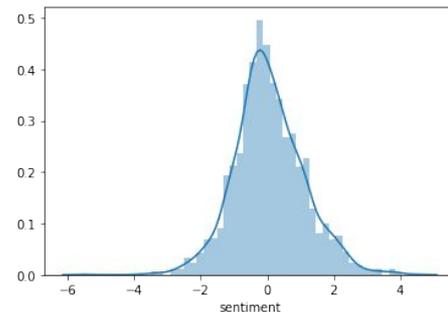


Medición
percepción
seguridad

<i>Senticon</i>	
<i>words</i>	<i>Sentiment</i>
w_1	$\rightarrow 0.2$
w_2	$\rightarrow 0.9$
\vdots	\vdots
w_n	$\rightarrow 0.5$

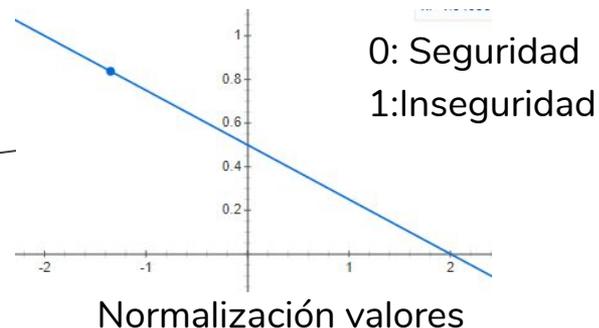
Cuantificación de los textos

$$\sum_{w \in W} \frac{\sum_{i \in P_w} S(i) + \sum_{j \in N_w} S(j)}{|P_w| + |N_w|}$$



Luz Rodríguez, 44 años de edad, Villavicencio. Fue atacada con arma de fuego en su casa por su expareja, quien irrumpió en ella violentamente. Su hija, en un intento de defenderla, también murió. Feminicida capturado.\n\nDiciembre, 2019 #2MNosotrasHablamos	0.886906
Refuerzan la seguridad en Bogotá por amenazas del ELN https://t.co/lQ1ru8229J	0.556109
En el peaje Andes ubicado a la salida norte de Bogotá, el ejército colombiano fortalece las medidas de seguridad para las personas y vehículos que quieran salir de la capital del país como medida de prevención para ... https://t.co/iMKmIV8sOt	0.050825
Aunque el hurto disminuyó, los homicidios aumentaron después de un mes de alivio. #ELTIEMPO habló con @HugoASeguridad, el secretario de Seguridad de Bogotá, y esto fue lo que nos dijo al respecto \n► https://t.co/rnGBk2jHWg https://t.co/sWlu8ZnBZZ	0.745832

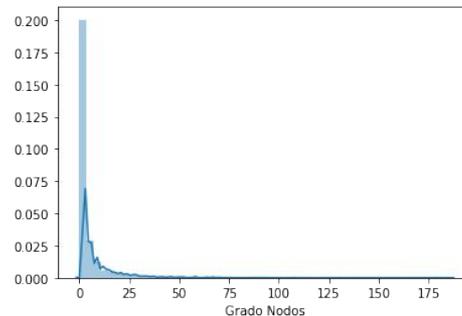
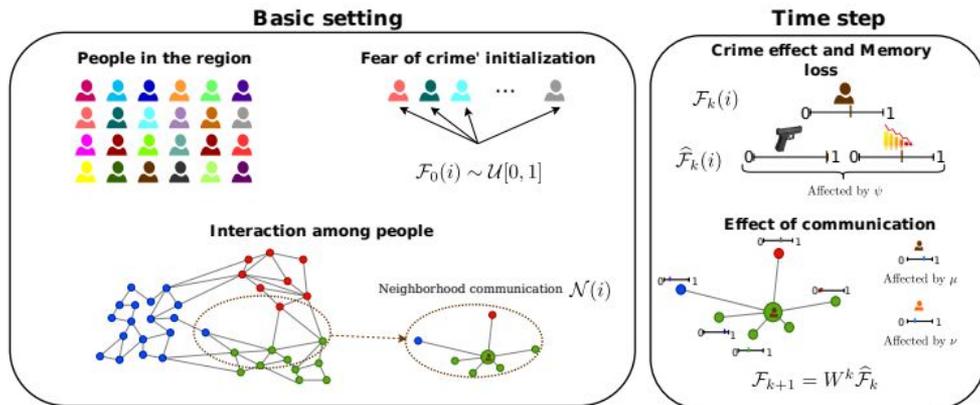
Ejemplos



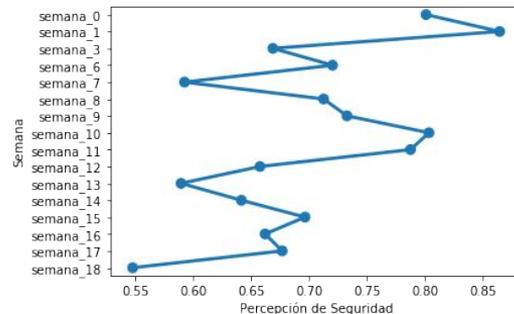
Normalización valores

Fenómeno de Percepción de Seguridad

Aplicación modelo de propagación de opinión



Distribución grado nodos grafo



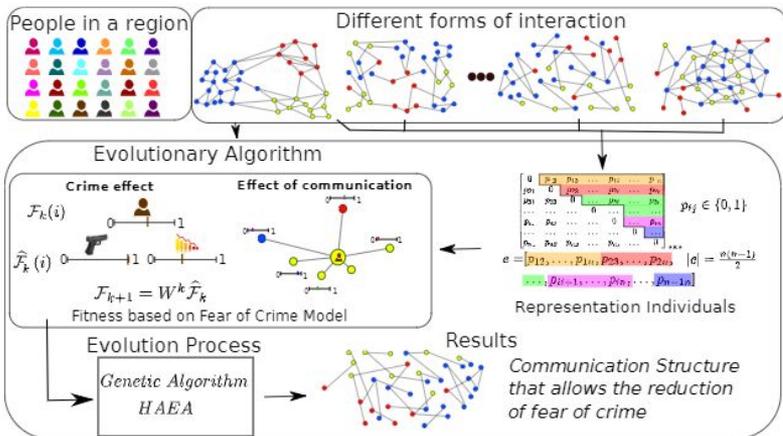
Evaluación media semanal



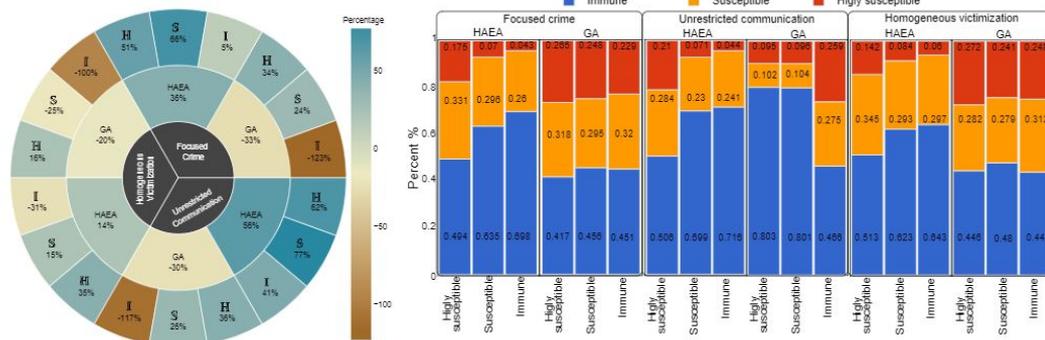
Fenómeno de Percepción de Seguridad

Pregunta de investigación: ¿Cuáles son las características de una estructura de comunicación que permite la comprender la propagación de los niveles de miedo al crimen en la ciudad de Bogotá?

Método



Resultados



La cohesión social es clave para mantener los niveles de miedo al crimen estables. Sin causar mayores riesgos a personas mayor afectadas con la victimización.



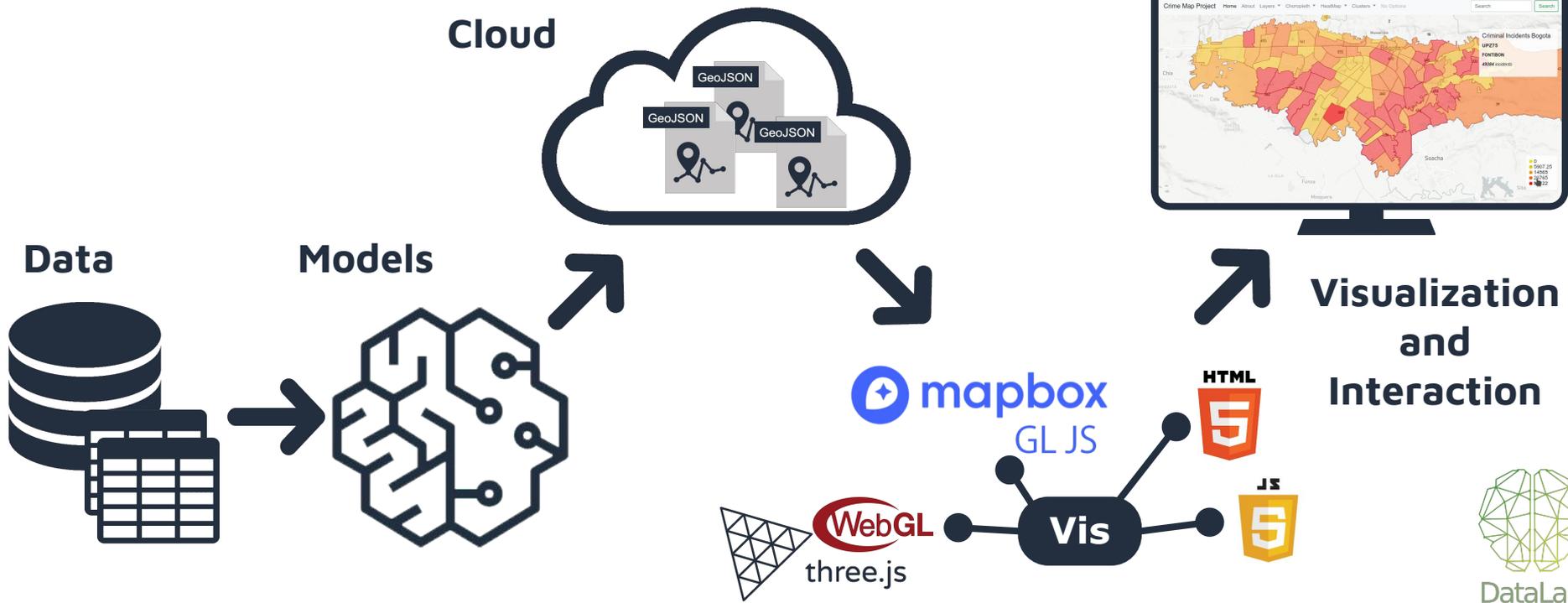
Herramienta de visualización interactiva de información del crimen



Sistema de visualización interactiva

Integración de los resultados de los modelos con el sistema

Pregunta de investigación: ¿Cómo se puede mejorar la interpretabilidad de los resultados mediante la visualización interactiva?



Estructura de datos de entrada

Mosaicos vectoriales

Formato: GeoJSON → Capas interactivas

- Geometrías: [vértices_2D (wgs84)]

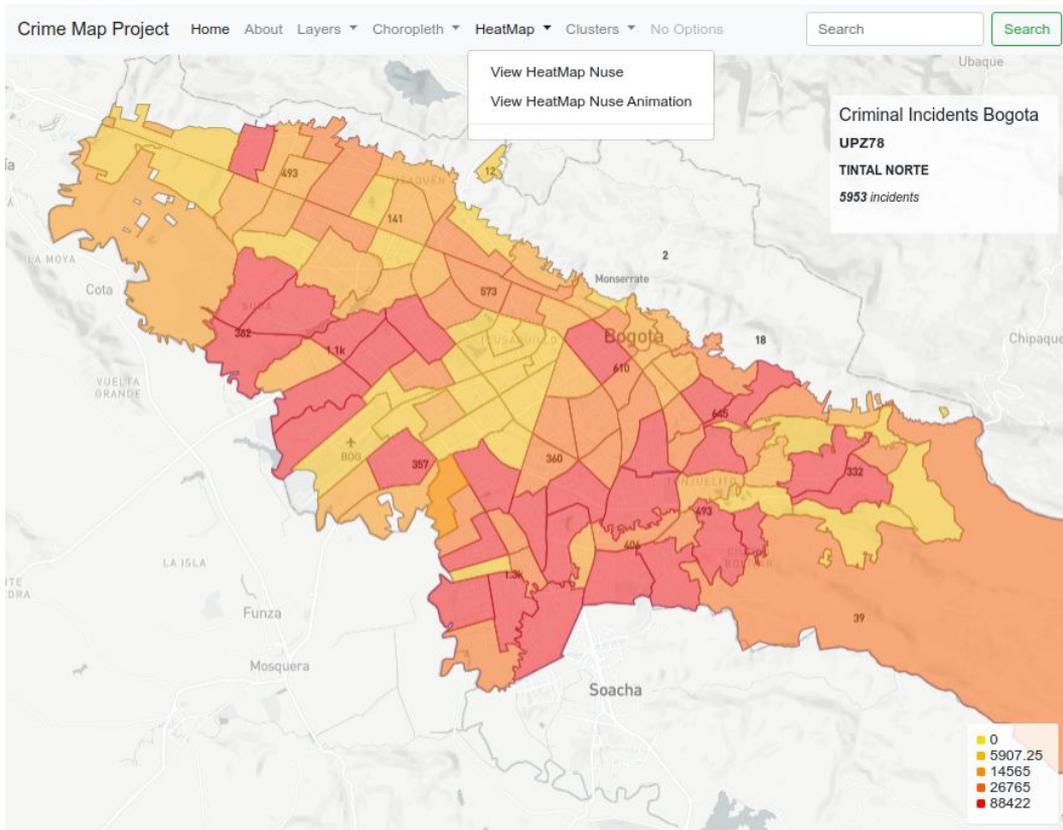
geometrías primitivas

Tipo	Ejemplos
Punto	
LineString	
Polígono	
	

geometrías compuestas

Tipo	Ejemplos
MultiPoint	
MultiLineString	
MultiPolygon	
	

- Atributos: [nombres, etiquetas, valores, pesos, categorías, listas]



Renderización de los datos

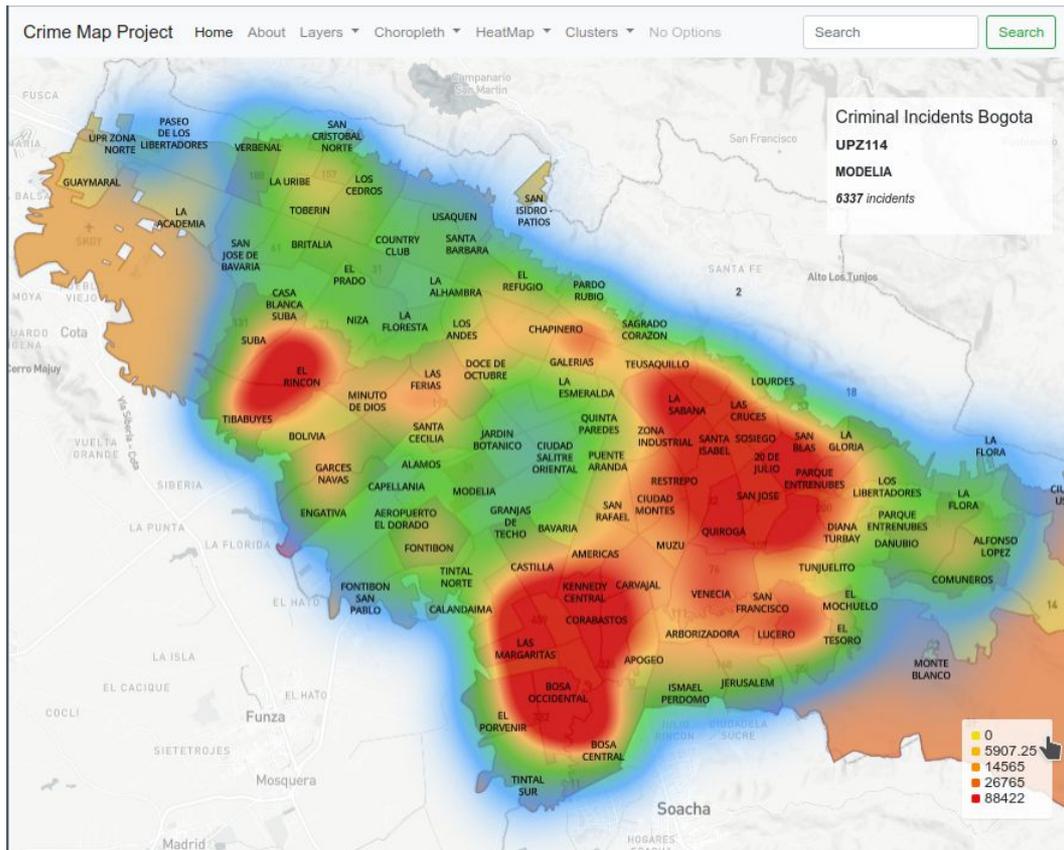
Estilos.mapboxGL-js

Estilos-capas

- Line
- Fill
- Circle
- Symbol
- Extrusion
- Three.js
- Heatmaps
- Iso-lines
- Iso-surfaces

Enriquecimiento

- Animación
- Agrupación de objetos
- Control de la escala
- Menús
- Elementos HTML5
- Eventos y presentación: bootstrap



GRACIAS

