

# Visualización dinámica de los crímenes en la ciudad de Chicago

Kelvin Paul Pucho Zevallos <sup>(1)</sup> Fabian Vladimir Florez Aguilar <sup>(2)</sup>

<sup>(1)</sup> Universidad Nacional de San Agustín, [kpucho@unsa.edu.pe](mailto:kpucho@unsa.edu.pe)

<sup>(2)</sup> Universidad Nacional de San Agustín, [fflorezag@unsa.edu.pe](mailto:fflorezag@unsa.edu.pe)

## 1 Introducción

La seguridad pública se ha convertido en un tema de debate constante, en el que la población y autoridades buscan estar más informadas sobre la incidencia delictiva en distintas áreas de la ciudad [1].

La delincuencia es un problema que afecta y preocupa a centros urbanos de todo el mundo, y la ciudad de Chicago no es una excepción, se ha observado que la población ha enfrentado a menudo diversas formas de actividad delictiva de diversos tipos, siendo la más frecuente el hurto. Además, ciertos sectores de la ciudad han experimentado un incremento drástico de delitos durante ciertos eventos o épocas del año. Como respuesta a esta problemática, la ciudad de Chicago ha estado recolectando datos sobre hechos criminales de diferentes tipos en diferentes áreas y distritos de la ciudad. La data que ellos proporcionan esta delimitada por diversos areas geograficas, ademas esta incluye informes policiales y registros públicos [2]. Chicago pone a disposición publica los registros de crímenes que sufrieron los ciudadanos, esto con el fin de comprender y estudiar el comportamiento de la delincuencia. Esto con el fin de poder analizar tendencias, identificar zonas con altos índices de delincuencia y descubrir patrones en la distribución de la delincuencia a lo largo del tiempo [3]. Con el fin de tener un mejor conocimiento de los crímenes se puede emplear diversas técnicas de análisis entre ellas estadística descriptiva, análisis de series temporales, análisis espacial, análisis de hotspot entre otras tantas técnicas. Estas técnicas puede ser usadas para resolver el problema de sesgos que presenta el dataset de Chicago, pues este surge por diferentes practicas policiales , dando como resultado una asignación desproporcionada de recursos policiales en zonas especificas.

Además la gran cantidad de datos proporcionados de Chicago plantea diversos desafíos en términos del análisis de datos espacio-temporal así como escoger la mejor representación visual además de aplicar un preprocesamiento a los datos.

En este contexto, se propone desarrollar un herramienta visual llamado “CrimeVis”, esta herramienta presenta diferentes visualizaciones del dataset de chicago crime esto con el objetivo de permitir a los autoridades y/o usuarios asignar recursos de forma proporcionada a la intensidad de delitos, esto apartir de observar y analizar diferentes visualizaciones de los delitos y con ello podra identificar diferentes patrones.

Ante ello se presenta las siguientes aportaciones:

1. Identificación de patrones cíclicos de los delitos mediante una visualización de calendario. Los datos de los crímenes se representarán de manera gráfica en un calendario, donde cada día del mes estará codificado con colores según la frecuencia de delitos ocurridos meses. El objetivo es permitir a los usuarios detectar patrones estacionales o cíclicos en la incidencia delictiva, identificando días o períodos con mayor actividad criminal recurrente a lo largo de un intervalo largo de tiempo.
2. Un análisis visual de los crímenes usando hotspot, para ello se divide el mapa en diferentes zonas geográficas (beat, wards,

community area y district). Esta visualización proporciona información sobre la incidencia, tipos de delitos y la densidad delictiva en diferentes zonas geográficas de la ciudad. Esto tiene como objetivo identificar patrones espaciales dentro de la ciudad de Chicago a paritr de ello las autoridades correspondientes puedan realizar un asignación de recursos proporcional a la intensidad de delitos.

3. Facilitar la identificación y análisis de patrones periódicos mediante visualizaciones de series temporales. Una visualización global de los delitos ayuda a las personas comprender la intensidad delictiva, sin embargo un análisis mas detallado de los crímenes en general y delitos específicos, ayudara a comprender como estos fueron evolucionando en el tiempo. El mapa esta dividido en sectores espaciales, donde cada sector mostrará las incidencias delictivas en diferentes intervalos de tiempo, estas están representadas mediante colores distintos que representa un crimen en especifico. Con ello podrán identificar los crímenes mas relevantes que se presentan en fechas especificas, con el cual se podría tomar medidas preventivas .

Este documento está organizado como sigue: La Sección 2 revisa trabajos relacionados, la Sección describe el marco teórico la Sección 4 describe la fuente de datos y el preprocesamiento del conjunto de datos, y la Sección 5 es una visión general del sistema. A continuación, en la Sección 6, presentamos en detalle las técnicas adaptativas de visualización de series temporales. En la Sección 7, presentamos casos de uso para demostrar la utilidad del sistema.

## 2 Trabajos Relacionados

En esta sección, llevamos a cabo una revisión de la literatura relacionada con las herramientas, técnicas y análisis de visualización para datos espaciales, temporales y espacio temporal. Se exploraron estudios clave que han investigado diversas técnicas de visualización utilizadas para representar datos complejos.

### 2.1 NOMBRE DEL PAPER

StanceXplore, este artículo propone una nueva herramienta de visualización NOMBRE DE LA HERRAMIENTA para la exploración interactiva de los post en los medios sociales. El objetivo de NOMBRE DE LA HERRAMIENTA es ayudar a los investigadores a explorar corpus de texto clasificados por post desde diferentes perspectivas simultáneamente, utilizando vistas múltiples coordinadas que incluyen temas definidos por el usuario, similitud y disimilitud de contenidos, y distribución geográfica y temporal. El artículo aborda el problema de explorar grandes volúmenes de datos textuales dinámicos y ruidosos procedentes de las redes sociales, lo que suele resultar difícil para las técnicas tradicionales de aprendizaje automático. Con NOMBRE DE LA HERRAMIENTA intenta ayudar a los investigadores de DH (QUE ES DH) la posibilidad de investigar patrones y

tendencias multidimensionales en la toma de postura relacionados con acontecimientos culturales, su distribución geográfica y la confianza del clasificador de posturas. El artículo utiliza un estudio de caso de usuarios de Twitter en Suecia para ilustrar cómo StanceX-plore puede utilizarse eficazmente para analizar el comportamiento de los usuarios de Twitter en cuanto a los temas tratados y las post adoptadas[4].



Figura 1: no mas otakus

## 2.2 NOMBRE DEL PAPER

Se presenta un sistema de análisis visual denominado UrbViz propone explorar la relación entre los índices urbanos y las características geográficas. UrbViz integra tres técnicas de visualización: clasificación multiatributo, análisis de puntos calientes y visualización de redes de causalidad. Estas técnicas están muy interrelacionadas y permiten a los usuarios analizar datos urbanos de forma exhaustiva e intuitiva. [5].



Figura 2: no mas otakus

## 2.3 Vismate: Interactive visual analysis of station-based observation data on climate changes

Vismate presenta un análisis y comprensión de los patrones de cambio climático 7 a diferentes escalas espacio- temporales. Reconoce que el cambio climático tiene un impacto significativo en diversos aspectos de la vida y se ha producido una serie de eventos de degradación ambiental relacionados con el cambio climático [6].

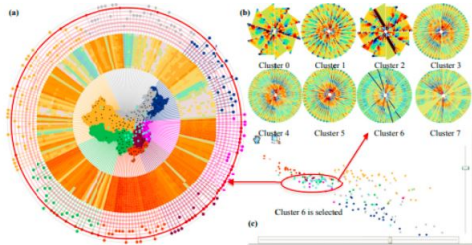


Figura 3: Vismate Cluster

## 2.4 A Visual Analytics Approach for Station-Based Air Quality Data

Presentan una herramienta llamada AirViz [7] que proporciona diversas herramientas de interacción para combinar la visualización

basada en mapas, vista de tendencias, vista de calendario y vista multidimensional. Además, presenta un controlador autoadaptativo basado en el calendario que puede adaptar con flexibilidad los cambios de tamaño y granularidad de los datos a la vista de tendencias 8 .

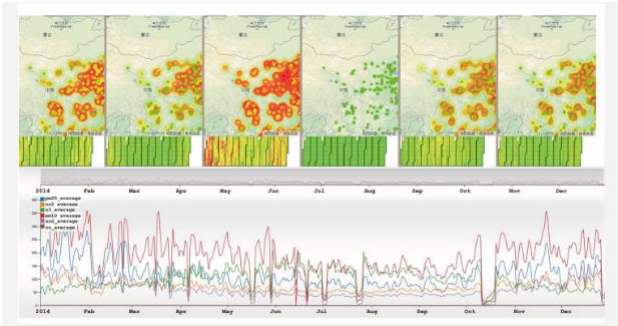


Figura 4: Vista multidimensional AirViz

## 2.5 CriPAV: Street-Level Crime Patterns Analysis and Visualization

Cripav presenta un enfoque consta de dos componentes principales. El primer componente utiliza un mecanismo estocástico para analizar visualmente los probables hotspot de delincuencia, revelando patrones importantes que no pueden observarse en la visualización típica basada en la intensidad de hotspot 5. El segundo componente utiliza un mecanismo de aprendizaje profundo para incrustar las series temporales de delitos en el espacio cartesiano, lo que permite identificar ubicaciones espaciales con comportamientos similares en las series temporales de delitos. Un nuevo método para identificar puntos conflictivos basados no sólo en el número de delitos, sino también en la probabilidad de que se produzcan. Combinando la probabilidad y la intensidad de los delitos en un gráfico de dispersión, nuestra metodología permite identificar visualmente los lugares donde se concentran los delitos, así como los sitios donde los delitos son frecuentes pero no necesariamente se producen en grandes cantidades [8].

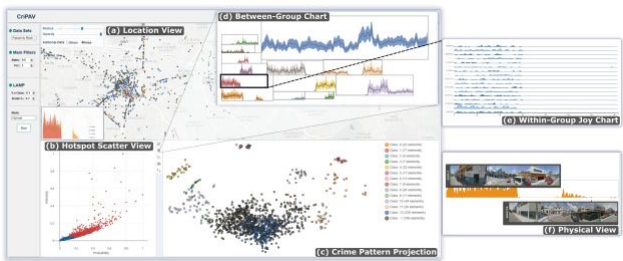


Figura 5: Cripav herramienta

## 2.6 NOMBRE DEL PAPER

El siguiente artículo presenta un nuevo enfoque para la visualización de datos de series temporales basado en espirales.[9] El problema que resuelve es la necesidad de una única técnica de visualización que sea apropiada para datos nominales, ordinales y cuantitativos, y que soporte la visualización de grandes conjuntos de datos. Los objetivos del artículo son presentar el gráfico en espiral como técnica de visualización para datos de series temporales, proporcionar técnicas para la visualización de datos de series temporales con

comportamiento periódico utilizando gráficos en espiral y ampliar el uso de espirales para la visualización de series temporales. Los autores también pretenden mejorar la identificación de ciclos en datos de series temporales y proporcionar herramientas adicionales para este fin.[9]



Figura 6: no mas otakus

## 2.7 NOMBRE DEL PAPER

El siguiente artículo es mas que todo visualizar patrones periódicos y presenta un nuevo método para explorar y analizar datos de series temporales extensas. [10] El problema que se aborda es cómo identificar patrones y tendencias en múltiples escalas temporales (días, semanas, estaciones) simultáneamente. La solución presentada consiste en agrupar patrones de datos diarios similares y visualizar los patrones medios en forma de gráficos y los días correspondientes en un calendario. Esta presentación proporciona una visión rápida tanto de los patrones estándar como de los excepcionales. Los objetivos de la ponencia son desarrollar métodos, técnicas y herramientas que permitan a los clientes analizar sus patrones de consumo energético de forma fácil y eficaz. El documento también pretende ofrecer un enfoque ascendente en el que el usuario pueda mostrar los clusters de arriba abajo y seleccionar el número de clusters que se mostrarán, a partir de lo cual el sistema genera una partición del año.



Figura 7: no mas otakus

## 2.8 NOMBRE DEL PAPER

En este artículo se analiza el problema de la extracción de patrones periódicos en datos de series temporales y su importancia práctica para mejorar la predicción y el análisis de tendencias. [11] El objetivo del artículo es clasificar los algoritmos actuales de extracción de patrones periódicos en tres categorías en función de si se conocen o no los parámetros de longitud periódica. El artículo también resume las ventajas y desventajas de estos algoritmos para proporcionar una referencia para su aplicación práctica. En resumen, el artículo pretende ofrecer una visión de la clasificación y evaluación de algoritmos para la extracción de patrones periódicos en datos de series temporales.

## 3 Tareas

- Visualizar las incidencias generales de los delitos y los crímenes más frecuentes a nivel de diferentes zonas geográficas (beats, wards, community area y district).

- Identificar patrones temporales. Permite identificar patrones mediante la visualización de series temporales por delitos.
- Explorar la vista de calendario para observar la distribución temporal de los datos de crímenes en diferentes periodos de tiempo, como días, semanas, meses y años. El usuario puede seleccionar un periodo de tiempo específico para ver la incidencia delictiva en ese intervalo, lo que permitirá identificar patrones, tendencias y variaciones en la actividad criminal a lo largo del tiempo.
- Los usuarios podrán elegir entre ver la frecuencia general de todos los delitos en Chicago a lo largo del tiempo o seleccionar delitos específicos para visualizar su frecuencia individual. Esto permitirá analizar y comparar la tendencia de diferentes tipos de delitos.

## 4 Marco Teórico

### 4.1 Pipeline

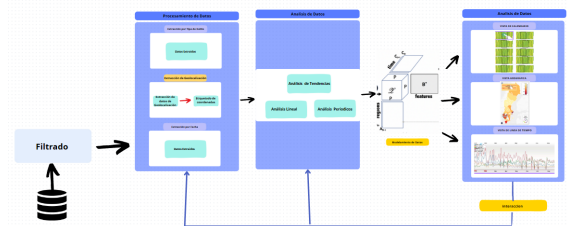


Figura 8: Vista multidimensional CrimeVis

### 4.2 Series Temporales

Las series temporales se refieren a un tipo de datos recopilados y organizados en orden cronológico, normalmente a intervalos equidistantes [12].

El análisis visual de series temporales consiste en analizar y visualizar los patrones, tendencias y características de los datos para obtener información y tomar decisiones fundamentadas [13].

Los datos de series temporales se utilizan habitualmente en campos como las finanzas, la economía, la predicción meteorológica y el procesamiento de señales, entre otro.

Una serie temporal es una secuencia de puntos de datos indexados o enumerados en orden temporal, donde cada punto de datos representa una observación en un momento específico [14]. Se trata de una colección de observaciones de un mismo sujeto reunidas a lo largo de diferentes intervalos de tiempo

Los datos de series temporales se organizan cronológicamente, siendo el tiempo el atributo central que los distingue de otros tipos de datos. Los puntos de datos se registran a intervalos constantes durante un periodo de tiempo determinado [15]. En cuanto al análisis visual de series temporales implica la creación de representaciones visuales, como diagramas de líneas o gráficos, para explorar y comprender los patrones y tendencias de los datos. Estas visualizaciones pueden ayudar a identificar patrones estacionales, tendencias, ciclos y otras características de los datos. [15] En el análisis visual de series temporales, pueden realizarse varias tareas de análisis, entre ellas: Exploración de motivos/formas: Identificación de patrones o formas recurrentes dentro de los datos de series temporales [13].



Identificación de patrones periódicos/cíclicos: Detección de patrones que se repiten a intervalos regulares, como ciclos diarios, semanales o anuales [13]. Identificación de tendencias: Analizar la dirección general y la magnitud del cambio en los datos a lo largo del tiempo [13]. En general, el análisis de series temporales en un contexto de análisis visual implica examinar los patrones y tendencias de los datos a través de visualizaciones para obtener información y tomar decisiones basadas en datos. Es una técnica valiosa para comprender el comportamiento y las características de los datos que dependen del tiempo.

### 4.3 Granularidad

La granularidad en un contexto de análisis visual se refiere al nivel de detalle o el grado de agregación en los datos que se analizan. Es el tamaño de la unidad de datos más pequeña que se está visualizando o analizando [16].

La granularidad de los datos puede afectar la precisión y la utilidad de las visualizaciones, así como el rendimiento del análisis. Por ejemplo, si los datos son demasiado granulares, puede ser difícil ver patrones o tendencias en los datos, mientras que si los datos están demasiado agregados, se pueden perder detalles importantes. [17]. La granularidad de una fórmula debe coincidir con la granularidad de la celda en el intervalo establecido antes de completar el cálculo. Si la celda está por debajo de la granularidad admitida, el resultado debe ser un espacio en blanco.

En el contexto de crímenes de Chicago, la granularidad puede afectar en la toma de decisiones de las autoridades en la asignación de recursos, definir el horario de los patrulleros o aplicar medidas preventivas durante los periodos de alto riesgo.

En general, la granularidad es un concepto importante en el análisis y visualización de datos, y es importante considerar detenidamente el nivel de detalle o agregación necesario para un análisis o visualización determinados.

### 4.4 HotSpot

El análisis de los hotspots es uno de los recursos empleados por las agencias de seguridad pública para planificar las patrullas policiales y diseñar acciones preventivas [18]. Los métodos de detección de puntos calientes suelen tener en cuenta el número absoluto de sucesos delictivos en cada lugar específico, dejando de lado los lugares en los que los delitos son probables pero no se producen en grandes cantidades, principalmente en comparación con su entorno. Los lugares con alta probabilidad de delitos pueden ser más perjudiciales para la comunidad que los lugares en los que la ola de delitos se produce en un breve periodo de tiempo [19].

La discretización espacial más común es una cuadrícula regular con una granularidad de celdas que varía en función de la escala a la que debe ir desde decenas de metros hasta grandes áreas que abarcan barrios enteros. Sin embargo, los delitos se concentran sobre todo en lugares "micro" que son relativamente estables a lo largo del tiempo [20].

### 4.5 Espacio-Temporal

La mayoría de las técnicas se basan en cuadrículas regulares con datos sobre la delincuencia en celdas cuadriculadas, cada una de las cuales abarca cientos de metros cuadrados. Sin embargo, estudios recientes señalan la importancia de analizar micro lugares, ya que la delincuencia rara vez se concentra en regiones mayores que un

segmento de calle o una esquina [21, 22]. El entorno de los lugares crea una oportunidad para delinquir como por ejemplo un bar, un parque, etc [23].

Un aspecto importante del comportamiento espacial de los delitos es la influencia de las regiones vecinas entre sí. Existe una mayor probabilidad que las regiones cercanas lo suficientemente pequeñas tengan características muy similares, por lo que es probable que tengan oportunidades similares para cometer delitos. Desde el punto de vista de la modelización, también es significativo que la importancia de una característica determinada en la determinación del comportamiento delictivo no sea necesariamente la misma para todas las regiones. Algunas características pueden ser más significativas cuando se dan juntas, mientras que otras pueden reducir los efectos de las demás. Por ejemplo, la presencia de tiendas en una región podría provocar un aumento de la delincuencia mayor que en una región más modesta. [24]

## Referencias

- [1] Pawale, P., Bagal, S., Ajabe, S., y Shikalagar, K., "Geo statistical approach for crime hotspot detection and prediction," 2017.
- [2] Alqahtani, A., Garima, A., y Alaiad, A., "Crime analysis in chicago city," pp. 166–172, 2019, doi:10.1109/IACS.2019.8809142.
- [3] Portal, C. D., "Crimes - 2001 to present." <https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-Present/ijzp-q8t2>. 2023.
- [4] Martins, R. M., Simaki, V., Kucher, K., Paradis, C., Kerren, A., *et al.*, "StanceXplore: Visualization for the interactive exploration of stance in social media," 2017.
- [5] Melo, M. A., UrbViz: Visual Analysis of Urban Indices and Geographically Aware Causality. PhD thesis, 2022.
- [6] Li, J., Zhang, K., y Meng, Z.-P., "Vismate: Interactive visual analysis of station-based observation data on climate changes," en 2014 IEEE Conference on Visual Analytics Science and Technology (VAST), pp. 133–142, 2014, doi:10.1109/VAST.2014.7042489.
- [7] Du, Y., Ma, C., Wu, C., Xu, X., Guo, Y., Zhou, Y., y Li, J., "A visual analytics approach for station-based air quality data," Sensors, vol. 17, p. 30, 2016, doi:10.3390/s17010030.
- [8] Garcia Zanabria, G., Medeiros Raimundo, M., Poco, J., Nery, M., Silva, C., Abreu, S., y Nonato, L., "Cripav: Street-level crime patterns analysis and visualization," IEEE Transactions on Visualization and Computer Graphics, vol. PP, pp. 1–1, 2021, doi:10.1109/TVCG.2021.3111146.
- [9] Weber, M., Alexa, M., y Müller, W., "Visualizing time-series on spirals," 2001.
- [10] Van Wijk, J. J. y Van Selow, E. R., "Cluster and calendar based visualization of time series data," 1999.
- [11] "An analysis of algorithms for periodic pattern mining in time series," 2019.
- [12] Berg, C., "The complete guide to times series data." <https://www.clarify.io/learn/time-series-data>. 03 November 2021.
- [13] de Zurich, U., "Visual analysis of time series." <https://www.ifl.uzh.ch/en/ivda/research/VA-of-TS.html>.
- [14] Community, W., "Time series." [https://en.wikipedia.org/wiki/Time\\_series](https://en.wikipedia.org/wiki/Time_series).
- [15] Tableau, "Time series analysis: Definition, types, techniques, and when it's used." <https://www.tableau.com/learn/articles/time-series-analysis>. 03 November 2021.
- [16] Ferrari, A., "Working below a dax formula granularity." <https://www.sqlbi.com/articles/working-below-a-dax-formulas-granularity/>. 17 August 2020.

- [17] Inmon, W. y Linstedt, D., “3.4 - data marts,” en *Data Architecture: a Primer for the Data Scientist* (Inmon, W. y Linstedt, D., eds.), pp. 115–119, Boston: Morgan Kaufmann, 2015, doi:<https://doi.org/10.1016/B978-0-12-802044-9.00018-0>.
- [18] Chainey, S., Thompson, L., y Uhlig, S., “The utility of hotspot mapping for predicting spatial patterns of crime,” *Security Journal*, vol. 21, pp. 4–28, 2008, doi:[10.1057/palgrave.sj.8350066](https://doi.org/10.1057/palgrave.sj.8350066).
- [19] Wang, D., Ding, W., Lo, H., Morabito, M., Chen, P., Salazar, J., y Stepinski, T., “Understanding the spatial distribution of crime based on its related variables using geospatial discriminative patterns,” *Computers, Environment and Urban Systems*, vol. 39, p. 93–106, 2013, doi:[10.1016/j.compenvurbsys.2013.01.008](https://doi.org/10.1016/j.compenvurbsys.2013.01.008).
- [20] Groff, E. R., Weisburd, D., y Yang, S.-M., “Is it important to examine crime trends at a local “micro” level?: A longitudinal analysis of street to street variability in crime trajectories,” *Journal of Quantitative Criminology*, vol. 26, pp. 7–32, 2010.
- [21] Weisburd, D., Bushway, S., Lum, C., y Yang, S.-M., *Trajectories of Crime at Places: A Longitudinal Study of Street Segments in the City of Seattle\**, pp. 443–476. 2017, doi:[10.4324/9781315089256-19](https://doi.org/10.4324/9781315089256-19).
- [22] Bernasco, W. y Block, R., “Robberies in chicago: A block-level analysis of the influence of crime generators, crime attractors, and offender anchor points,” *Journal of Research in Crime and Delinquency*, vol. 48, pp. 33–57, 2011, doi:[10.1177/0022427810384135](https://doi.org/10.1177/0022427810384135).
- [23] Wright, R., Decker, S. H., y Shover, N., “Armed robbers in action: Stickers and street culture,” 1997.
- [24] Garcia-Zanabria, G., Gomez-Nieto, E., Silveira, J., Poco, J., Nery, M., Adorno, S., y Nonato, L. G., “Mirante: A visualization tool for analyzing urban crimes,” en *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 148–155, 2020, doi:[10.1109/SIBGRAPI51738.2020.00028](https://doi.org/10.1109/SIBGRAPI51738.2020.00028).