



Visual Analytics de Datos Epidemiológicos: Análisis de la Propagación y Olas Pandémicas del COVID-19

CARAZAS QUISPE, Alessandro Jesus

Orientador:

**UNSA - Universidad Nacional de San Agustín de Arequipa
Junio de 2025**

1. Introducción

La pandemia de COVID-19 evidenció la importancia de contar con herramientas avanzadas para analizar datos epidemiológicos de forma eficiente y oportuna. La disponibilidad creciente de datos abiertos —que incluyen registros diarios de casos confirmados, muertes y localización geográfica— ha generado oportunidades para estudiar la propagación del virus y la aparición de olas epidémicas en distintas regiones del mundo. Este escenario plantea la necesidad de integrar enfoques computacionales y visuales que permitan no solo representar los datos, sino también interpretarlos en sus dimensiones espacio-temporales y multivariadas.

En particular, el concepto de propagación ha adquirido nuevas interpretaciones a raíz de estudios recientes sobre el comportamiento del SARS-CoV-2. Hernández-Vargas y Velasco-Hernández [1] plantean que la transmisión del COVID-19 no debe entenderse como un proceso lineal o pasivo, sino como un fenómeno *cooperativo*, en el cual múltiples exposiciones virales se acumulan hasta superar un umbral crítico que dispara abruptamente la probabilidad de infección. Esta idea introduce una dimensión no lineal a la propagación, donde ciertas condiciones —como la densidad poblacional, la movilidad o la ventilación deficiente— pueden acelerar súbitamente el ritmo de contagios. Comprender este tipo de dinámicas resulta esencial para analizar el surgimiento de brotes epidémicos intensificados y la evolución de olas pandémicas.

En este contexto, la *visual analytics*, disciplina que combina visualización interactiva con análisis de datos, se presenta como una herramienta valiosa para abordar la complejidad del fenómeno pandémico. Su aplicación permite explorar relaciones dinámicas entre variables clínicas, temporales y geográficas, favoreciendo la identificación de patrones como la aparición de nuevas olas epidémicas, la velocidad de propagación entre territorios y la magnitud del impacto sanitario en distintas regiones [2, 3].

Sin embargo, persisten desafíos relevantes. Muchas de las herramientas utilizadas durante la pandemia presentaron limitaciones al integrar múltiples escalas espaciales (país, región, continente), representar tendencias temporales prolongadas y manejar la incertidumbre inherente a los datos epidemiológicos [4, 5]. Además, la falta de interactividad en gran parte de estas soluciones dificultó el trabajo exploratorio por parte de epidemiólogos y tomadores de decisiones [6]. Esta situación generó una brecha entre la disponibilidad de los datos y su análisis efectivo, afectando la capacidad de respuesta oportuna en salud pública [7].

A partir de esta problemática, surge la motivación por desarrollar un enfoque integral que permita analizar la propagación del COVID-19 y caracterizar las olas pandémicas a través de visualizaciones dinámicas e interactivas. Este trabajo tiene como objetivo analizar datos epidemiológicos del COVID-19 mediante técnicas de *visual analytics*, con el propósito de identificar y caracterizar patrones de propagación y olas pandémicas, evaluando su utilidad para la comprensión del fenómeno y el apoyo a la toma de decisiones en salud pública.

2. Trabajos Relacionados

Diversos estudios han abordado la propagación del COVID-19 desde perspectivas interdisciplinarias, reconociendo la complejidad del fenómeno pandémico en relación con factores biológicos, sociales y espaciales. A continuación, se presentan dos trabajos que aportan fundamentos conceptuales clave para este proyecto.

El estudio de Hernandez-Vargas y Velasco-Hernandez titulado “*Cooperative virus propagation in COVID-19 transmission*” [1] plantea una aproximación novedosa al concepto de propagación viral. En lugar de considerar una transmisión lineal o aditiva, los autores proponen un mecanismo de propagación cooperativa, en el cual la probabilidad de infección aumenta de manera no lineal con la exposición viral acumulada. Esta característica implica que existe un umbral crítico de exposición que, una vez superado, incrementa drásticamente la probabilidad de infección. Esta perspectiva es relevante para la visualización de la propagación del COVID-19, ya que permite interpretar brotes epidémicos intensificados en zonas donde la densidad poblacional, la movilidad o la exposición acumulada favorecen estos “efectos umbral”. El estudio argumenta además que este tipo de propagación requiere intervenciones más intensas y rápidas para ser contenida, lo cual refuerza la necesidad de herramientas visuales que permitan detectar focos de alta acumulación viral y actuar en consecuencia.

Por otro lado, el trabajo de Healey et al. “*Visual Analytics for Understanding the Spread of COVID-19*” [8] explora cómo la combinación de análisis de datos multivariados y técnicas de visualización interactiva puede mejorar significativamente la comprensión del fenómeno de propagación global del virus. El enfoque se basa en representar datos espaciales y temporales mediante dashboards y mapas dinámicos, los cuales permiten identificar la aparición y evolución de brotes en diferentes regiones. Los autores destacan que el análisis visual facilita la detección de patrones de propagación complejos, como las olas epidémicas asincrónicas entre países, así como la correlación entre factores como densidad de población, movilidad y medidas sanitarias adoptadas. Este trabajo proporciona una base conceptual y técnica que respalda la propuesta de este proyecto, al mostrar el potencial del *visual analytics* para convertir grandes volúmenes de datos epidemiológicos en conocimiento accionable.

Ambos trabajos refuerzan la idea de que la propagación del COVID-19 no puede entenderse únicamente como un proceso lineal o homogéneo. Por el contrario, debe concebirse como un fenómeno dinámico, multiescalar y condicionado por diversos factores biológicos, sociales y geoespaciales. En este sentido, el estudio de la propagación y de las olas pandémicas del COVID-19 requiere enfoques capaces de integrar múltiples dimensiones temporales y espaciales de forma interactiva y comprensible. Bajo esta perspectiva, la presente investigación titulada *Visual Analytics de Datos Epidemiológicos: Análisis de la Propagación y Olas Pandémicas del COVID-19* propone un *framework* visual que permita explorar estos patrones complejos a través de herramientas interactivas, facilitando su interpretación por parte de epidemiólogos, analistas y tomadores de decisiones.

3. Propuesta

La presente propone **analizar datos epidemiológicos del COVID-19** mediante un *framework* de *Visual Analytics* que permita la exploración interactiva espacio-temporal de la propagación del virus. El enfoque se centrará en identificar patrones de transmisión, zonas críticas y la evolución temporal del brote a través de visualizaciones dinámicas y multiescales, sin utilizar modelos predictivos. La propuesta busca superar las limitaciones de las plataformas tradicionales mediante el uso de mapas geoespaciales interactivos, series temporales animadas y filtros exploratorios que faciliten la comprensión profunda del fenómeno pandémico. El análisis se apoyará en un backend eficiente en Python y una base de datos espacial en PostgreSQL/PostGIS, utilizando datos históricos del COVID-19 como caso de estudio.

3.1. 3.1 Pipeline Propuesto

El pipeline propuesto para el análisis espacio-temporal de los datos epidemiológicos del COVID-19 está diseñado para permitir una exploración interactiva y multiescalar de los brotes pandémicos. Cada etapa del pipeline contribuye a estructurar, procesar y visualizar la información de manera que los usuarios puedan identificar patrones de propagación, zonas críticas y tendencias temporales relevantes.

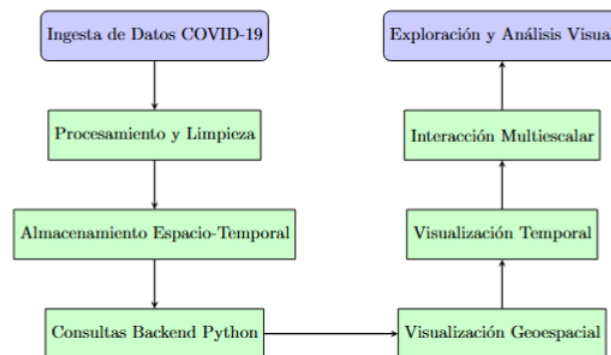


Figura 1: Pipeline propuesto para el análisis espacio-temporal de datos del COVID-19.

A continuación, se describen las etapas del pipeline y las visualizaciones que se pretenden lograr:

- 1. Recolección y Normalización de Datos:**

Se recopilarán datos históricos del COVID-19 provenientes de fuentes abiertas, asegurando su limpieza y estandarización para facilitar el análisis.

- 2. Almacenamiento Espacio-Temporal:**

Los datos serán estructurados y almacenados en una base de datos espacial utilizan-

do PostgreSQL y PostGIS, permitiendo consultas eficientes por región y por rango de tiempo.

3. **Procesamiento y Consulta:**

El backend gestionará la recuperación de datos filtrados por tiempo y ubicación, optimizando las consultas para soportar visualizaciones interactivas en tiempo real.

4. **Visualización Geoespacial:**

Se generarán mapas interactivos donde los usuarios podrán observar la distribución de casos por región, utilizando mapas de calor, mapas coropléticos y clústeres dinámicos.

5. **Visualización Temporal:**

Se implementarán series temporales animadas que permitirán explorar la evolución de los casos a lo largo del tiempo. Se incluirán sliders y animaciones que faciliten el análisis dinámico.

6. **Interacción Multiescalar:**

Los usuarios podrán aplicar filtros y explorar diferentes escalas geográficas (global, nacional, regional, local) y temporales (días, semanas, meses) de forma interactiva.

7. **Validación con Caso de Estudio:**

El pipeline será evaluado utilizando datos históricos del COVID-19, verificando su capacidad para facilitar la detección de patrones no evidentes mediante exploración visual.

Este pipeline busca lograr un sistema de visualización interactivo que permita a los usuarios comprender la propagación del COVID-19 desde múltiples perspectivas, facilitando la toma de decisiones informadas en salud pública.

Referencias

- [1] E. A. Hernández-Vargas and J. X. Velasco-Hernández, "Cooperative virus propagation in covid-19 transmission," *medRxiv*, 2020.
- [2] Y. Dong, C. J. Liang, Y. Chen, and J. Hua, "A visual modeling method for spatio-temporal and multidimensional features in epidemiological analysis: Applied covid-19 aggregated datasets," *Computational Visual Media*, vol. 10, no. 1, pp. 161–186, 2024.
- [3] D. Kim, B. Cánovas-Segura, M. Campos, and J. M. Juarez, "Visualization of spatial-temporal epidemiological data: A scoping review," *Technologies*, vol. 12, no. 3, p. 31, 2024.
- [4] T. Joliveau, R. Cura, and H. Commenges, "Covid-19 geoviz for spatio-temporal structures detection," in *Proceedings of the International Cartographic Association*, vol. 4, p. 37, 2021.

- [5] S. Luz and M. Masoodian, “Temporal and spatial elements in interactive epidemiological maps.” arXiv preprint arXiv:2206.06048, 2022.
- [6] University of Memphis School of Public Health, “Visual communication of public health data: A scoping review,” *Frontiers in Public Health*, vol. 13, p. 1234567, 2025.
- [7] J. Chen, A. M. MacEachren, and D. Guo, “Visual analytics for epidemiologists: Understanding the interactions between age, time, and disease with multi-panel graphs,” *PLOS ONE*, vol. 6, no. 2, p. e14683, 2011.
- [8] C. G. Healey, J. Smith, T. Wang, and A. Li, “Visual analytics for the coronavirus covid-19 pandemic,” *Big Data*, vol. 10, no. 2, pp. 123–135, 2022. Acceso libre.