

# Big data para el análisis de la movilidad en Bogotá

Camilo Andrés Nemocón Farfán  
Secretaría de Movilidad  
Bogotá, Colombia

**Resumen— Este artículo da a conocer el desarrollo de una plataforma web que permite realizar consultas de caracterización de corredores, aforos, velocidades, incidentes, siniestros, alertas de Waze, PMT y comparendos, siendo esta una herramienta para los gerentes de zona de la Secretaría de Movilidad de Bogotá, con la cual analizan la información histórica y geográfica para medir la evolución y el comportamiento de la movilidad desde distintas áreas, y así generar los indicadores de gestión de la ciudad de Bogotá en el sector de la movilidad.**

**Palabras Claves— Analítica visual, sistemas urbanos, movilidad, visualización, big data.**

## I. INTRODUCCION

La ciudad de Bogotá es una de las ciudades donde el uso del suelo se encuentra disperso y distante, generando unas necesidades e intensidades de movilización más altas de Colombia, debido a que las modificaciones en los usos de la tierra cambian en torno al crecimiento y flujo dinámico de la ciudad.

Por tanto, para un buen desarrollo y planeación de la movilidad en torno al crecimiento y comportamiento actual de la ciudad y su movilidad, se buscan crear condiciones viables que permitan el desplazamiento a partir de la demanda donde prima el interés general sobre el particular, restándole protagonismo a las políticas basadas en la provisión de la infraestructura y enfocándose más en el comportamiento de los ciudadanos [1].

Para tomar las mejores decisiones y planeaciones en torno a la demanda de la movilidad, se necesita que los expertos tengan acceso del comportamiento dinámico de desplazamiento de los ciudadanos, los cuales se pueden obtener a partir de datos recolectados de distintas fuentes y almacenados en una granja de servidores en la Secretaría de Movilidad de Bogotá. Por otro lado, no basta con guardar la información sino poderla manejar, procesar, explorar y visualizar simultáneamente, para descubrir patrones, relaciones, tendencias u otras características que apoyen al proceso de análisis, planteamiento de hipótesis y toma de decisiones que realizan los expertos.

## II. PROBLEMA Y SOLUCIÓN

Actualmente la Secretaría de Movilidad de Bogotá (SDM) cuenta con 3 herramientas para la visualización y manejo de datos, estas herramientas son la Tableau, ArcGis y OracleBI.

Sin embargo, ninguna de estas herramientas están integradas, es decir, aunque la información está concentrada, no existe una plataforma que unifique la información entorno al análisis de los datos, por ejemplo, utilizan ArcGis para mostrar todos los datos georreferenciados pero no se pueden realizar cálculos, promedios, tabulaciones o gráficos de utilidad [2] que ayuden a un análisis sin necesidad de pasar la información visualizada a un Excel. Por otro lado, se tiene Oracle BI, la cual es una herramienta que permite procesar y mostrar gráficos de utilidad a partir de los datos, sin embargo, este framework es rígido y no le brinda al usuario la libertad de generar modificaciones sobre las gráficas, consultas o filtros de los datos, más allá de los predefinidos por la aplicación.

Finalmente se tiene Tableau, el cual cumple con todos los requerimientos para el proceso de analítica visual, teniendo en cuenta conceptos como el área de consulta o filtrado de la información, gráficos de utilidad y mapa focalizado [2], con lo cual permite una fácil exploración sobre los datos permitiendo visualizar los resultados de forma georreferenciada sobre un mapa y mostrar gráficos entorno a solo la información requerida por el usuario.

Por tanto, se desarrolló de una plataforma que integra los datos espacio - temporales, siguiendo un protocolo que permite determinar la calidad de los datos [3], para su posterior uso dentro de éste sistema la información mediante la herramienta de Tableau para la visualización de datos, ya que a partir de los datos se pueden tomar decisiones y realizar análisis por medio de la plataforma web vinculando la analítica visual y las consultas rápidas sobre la gran cantidad de datos que se maneja diariamente en la Secretaría de Movilidad Bogotá.

## III. METODOLOGÍA

En el área de movilidad no existe una metodología específica para el manejo y visualización de la información, por tanto, se utilizó la planteada por Ben Fry [4], donde se adquieren los datos, independiente de la fuente o la descripción de la información, para posteriormente estructurarlos y organizarlos, normalizando la información y

así poder realizar consultas de acuerdo a las preguntas que se desean responder los expertos o mostrar las historias que nos presentan los datos.

Finalmente, este proceso es iterativo, generando una representación visual de los datos mediante áreas de consulta, área de resultados como mapas y áreas de estadísticas [5], que van siendo modificados de acuerdo a las interacciones de usabilidad que el usuario necesite.

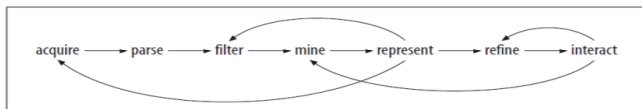


Figura 1. Metodología para el manejo y visualización de la información.

En paralelo a la anterior metodología, se manejó el modelo de la analítica visual generando una relación entre las visualizaciones presentadas por los datos y la cognición o entendimiento de la información [6], la cual se realiza mediante la interacción con la plataforma que responde a las consultas de datos y al procesamiento de los mismos mediante un modelo, para así mostrar la información para el análisis, resolución de problemas y toma de decisiones por parte del usuario.

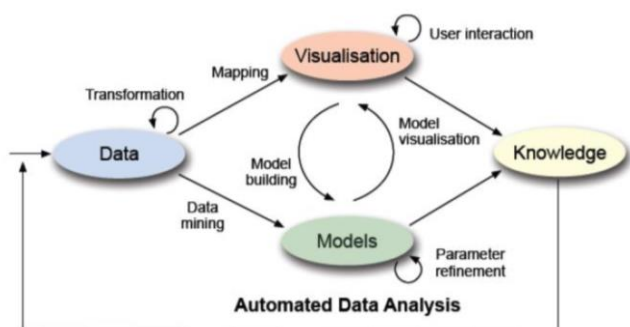


Figura 2. Modelo de la analítica visual.

Teniendo en cuenta estas metodologías, se procedió a recopilar los datos entorno a las áreas de movilidad como la seguridad vial, gestión, logística, tecnología e infraestructura, sobre las cuales se tiene un gran volumen de datos y que permiten determinar y analizar la velocidad por corredores y por tiempos, así mismo, visualizar los siniestros, incidentes y alertas, volúmenes y los comparendos.

Una vez recopilada la información se generan vistas de datos organizados, normalizando toda la información en tres categorías: espacio donde toda información debe poseer latitud, longitud, nombre o dirección, la otra categoría es el tiempo y su granularidad [3], años, meses, semanas, días, horas. La última categoría es el atributo, dado por el valor del tipo de dato, es decir, el valor de la velocidad, o el tiempo de viaje, el tipo de siniestro, el actor vial, el tipo de alerta, etc.

Se genera la conexión a la base de datos y se extrae la data para no manejarla en tiempo real y no ocupar una conexión permanente a la base de datos, en vez de eso, se extrae la

información semanal y se carga en la plataforma guardando todo el histórico.

Posteriormente se desarrollan tableros o dashboards en Tableau, Google Maps y HTML5, para representar y visualizar los datos, generando controles de interacción para realizar consultas y filtros sobre la información entorno a la formulación de una hipótesis o a una pregunta a resolver.

#### IV. DISEÑO E IMPLEMENTACIÓN

Esta plataforma integra la información histórica de velocidades, aforos, caracterización de corredores, incidentes del CGT (Centro de Gestión de Tránsito), siniestros SIGAT (Sistema de Información Georeferenciado de Accidentes de Tránsito), alertas de Waze, PMT (Plan de Manejo del Tráfico) programados y comparendos impuestos, generando relaciones entre los datos de distintas áreas para así identificar patrones de comportamiento de los ciudadanos a partir de las ubicaciones, usos del suelo, fechas y días.

Los usuarios de esta plataforma son los gerentes de área de la Dirección de Control y Vigilancia, los cuales son ingenieros que se encargan de hacerle seguimiento a los principales corredores viales de su área y proponen soluciones a los puntos críticos detectados. Estas pueden ser cambios en la señalización vial para organizar mejor los flujos vehiculares, reprogramación de semáforos, solicitar apoyo del grupo operativo en vía, o la ejecución de operativos con Policía de Tránsito, entre otras.

Por tanto, esta plataforma está desarrollada y dividida en diferentes secciones, a partir de las necesidades de uso de los gerentes de área, donde en cada sección se visualiza los datos que permiten medir y generar indicadores en la movilidad de Bogotá.

La primera sección es la de inicio en donde se caracterizan todos los corredores de Bogotá, es decir, se ubican y se pintan sobre el mapa cada uno de los corredores y éstos son divididos en los tramos determinados por cada gerente de zona, mostrando información relevante del territorio sobre el cual se encuentra cada tramo como: el número de población, estrato, calzadas, carriles, ciclorutas, capacidad en la vía, restricciones de carga, cantidad de intersecciones, centros atractores, puntos críticos, nombres de los barrios y localidades por donde pasa cada corredor. Esto permite comprender el valor estratégico de cada corredor en la movilidad de la ciudad, con respecto al número de personas que impacta. Por lo tanto, permite priorizar las actuaciones de la Secretaría para mejorar la movilidad de los bogotanos.



Figura 3. Plataforma SDM, sección de inicio.  
<http://201.245.176.195:8083/SDMplataforma/index.html>

La siguiente sección consiste en los aforos, en donde se muestra la información del conteo de vehículos desde el año 2011 en adelante, mostrando el resumen del aforo indicando el total de vehículos por tipo en cada una de las intersecciones aforadas. Esta sección permite que el experto pueda analizar la evolución del tránsito en sus corredores a lo largo de los años, y comprender cuales son las horas de máxima demanda para proponer medidas de gestión de tránsito.

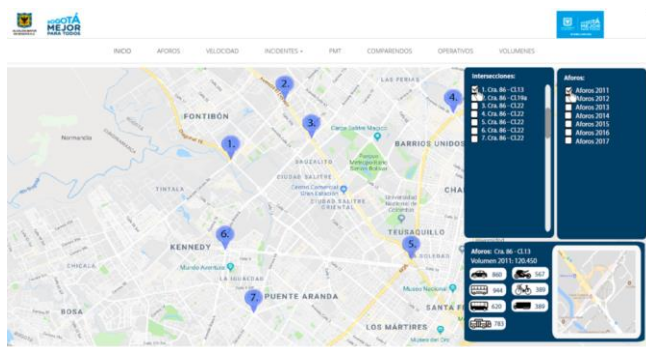


Figura 4. Plataforma SDM, sección de aforos.  
<http://201.245.176.195:8083/SDMplataforma/aforos.html>

En otra sección se encuentran los datos correspondientes a las velocidades en donde se evidencia el comportamiento de la velocidad promedio hora a hora, a partir de la selección de filtros de tiempo (Años, mes, día, hora) y de espacio, (corredor, tramos), de esta forma se puede ver la velocidad durante las 24 horas de un corredor en particular o de un conjunto de corredores.

Estos datos se obtienen principalmente de 2 fuentes de información, la primera es Bitcarrier el cual es un sistema de información de velocidad en corredores en tiempo real que consta de 350 sensores de wifi y bluetooth, 12 de conteo de bicicletas y 4 de volúmenes vehiculares en la ciudad de Bogotá. Estos sensores captan las señales públicas de wifi y bluetooth de los dispositivos móviles de los usuarios (celulares, radios bluetooth de automóviles, GPS, etc.) y envían anónimamente la información a una base de datos donde la data se procesa y se realizan todos los cálculos sobre velocidades y tiempos de recorrido que se muestran posteriormente en una interfaz gráfica.

La segunda fuente de información es Waze, el cual es una red social de movilidad alimentada por los usuarios del App.

Sin embargo, la Secretaría de Movilidad de Bogotá tiene un convenio con Waze, por tanto, tiene acceso a una plataforma llamada “Traffic View” con la cual se pueden dibujar hasta 199 tramos y obtener la velocidad en esos puntos de la ciudad.

A partir del desarrollo de éste tablero se puede calcular la velocidad promedio de Bogotá, teniendo en cuenta que es una velocidad calculada, es decir, que aparte de tomar el dato de la velocidad, también se tiene en cuenta la longitud de los tramos y corredores para sacar una velocidad unificada para Bogotá.

Por otro lado, esta sección permite comparar la velocidad entre años y por meses, con lo cual se puede confrontar distintos años, meses y días, evidenciando las mejoras que se han ido ejerciendo mediante las políticas de movilidad implementadas.



Figura 5. Plataforma SDM, sección de velocidad.  
<http://201.245.176.195:8083/SDMplataforma/velBogota.html>

Para entrar en la sección de incidentes del Centro de Gestión de Tránsito, es relevante describir que estos datos corresponden a todos los incidentes que son reportados por los ciudadanos vía Twitter y mediante la app Simur (Sistema de Información de Movilidad Urbana). A continuación, los incidentes son subidos a un servicio de información llamado “Incidents Reporting Platform” en donde se monta la información reportada.

En esta sección se muestra la cantidad de incidentes y los tipos de incidentes (accidente de tránsito simple, con heridos o muertos, inundación, semáforo dañado o vehículo varado). Estos datos son georreferenciados en el tablero a partir de la selección temporal (año, mes, día y hora) y la selección espacial (localidad), generando una visualización de múltiples mapas, donde la cantidad de mapas depende del número de días seleccionados, con lo cual se presentan los incidentes de cada día, permitiendo su comparación para así determinar patrones de ubicación.

Así mismo, se presentan gráficas estadísticas en donde se muestran la cantidad de incidentes por localidad, por hora, por tipo de incidente, permitiéndole a los expertos comparar y ver la cantidad y el tipo de incidente que más se generan por localidad y ubicación geográfica para así implementar una mejor asignación de los recursos espacialmente y permite identificar puntos críticos para hacer mayor control en estos puntos y disminuir la accidentalidad.

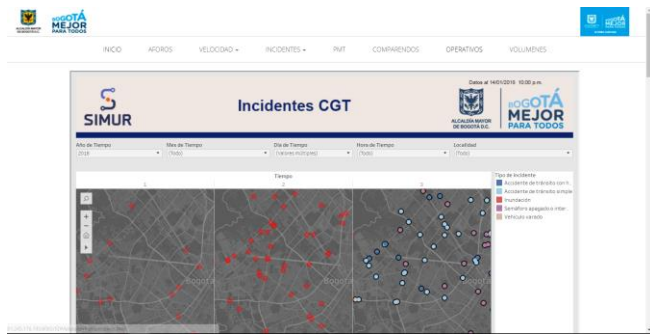


Figura 6. Plataforma SDM, sección de incidentes.  
<http://201.245.176.195:8083/SDMplataforma/incidentes.html>

Por otro lado, encontramos las alertas de Waze, es decir, se obtienen los datos de las alertas generadas por los usuarios de la App, como embotellamientos, siniestros, cierres viales, vehículos mal estacionados, árboles caídos, entre otras, y donde esta información se guarda desde abril del 2017.

Debido a que esta información se obtiene en tiempo real, se desarrolló un framework que permite visualizar los datos y así se pueden gestionar los recursos necesarios ante la ocurrencia de incidentes en la vía, sin embargo, es de tener en cuenta, que los datos deben ser pre-procesados para identificar y mostrar los datos que sean confiables, es decir, que las acciones de gestión de tránsito se hagan efectivas únicamente con las alertas que realmente sean ciertas y verídicas.

Por tanto, la confiabilidad de los datos de las alertas viene a partir de la cantidad de los “likes” de los usuarios frente a un mismo incidente, así mismo, como el puntaje o calificación que posee el usuario que genero la alerta y finalmente cruzando la información de todas las personas que publicaron el incidente en la misma ubicación.

La visualización de estos datos en la plataforma muestra la cantidad de alertas y los tipos de alertas a partir de la selección temporal (año, mes, día y hora) y la selección de atributos (tipo y subtipo de alerta), presentando un mapa, el cual se divide en la cantidad de días seleccionados, con lo cual se presentan las alertas de cada día, permitiendo su comparación para así determinar comportamientos georreferenciados.

En esta misma sección se presentan gráficas en donde se muestran la cantidad de alertas y subtipos de alertas por hora, lo cual les ayuda a los expertos a identificar y conocer los incidentes en la vía, para así gestionar los recursos y evitar el bloqueo en la vía en el menor tiempo posible.

Esta herramienta permite verificar los tipos de alertas y a partir de estos datos de geo referenciación y el histórico del tiempo en que se generan las alertas, se puede determinar si en un mismo punto se presentan alertas de siniestros, para evaluar la zona e identificar posibles causas de siniestros debido a la vía. Así mismo, permite analizar la periodicidad en que se presentan estos incidentes para llegar a determinar un patrón de ubicación, de días y de horas en que se generan más alertas y el tipo de alerta.

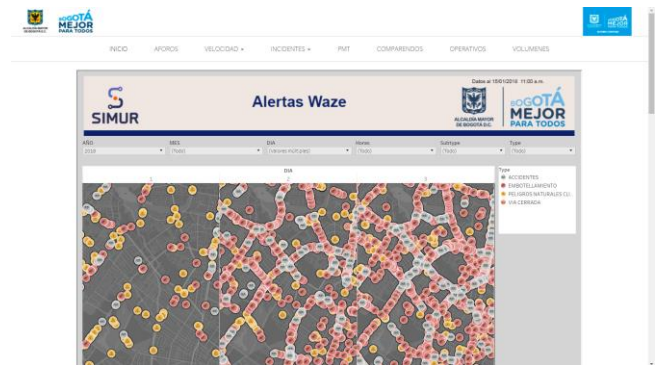


Figura 7. Plataforma SDM, sección de alertas.  
<http://201.245.176.195:8083/SDMplataforma/alertas.html>

Continuando con la sección de siniestralidad, los datos que se manejan corresponden a la base de datos SIGAT (Sistema de Información Georreferenciado de Accidentes de Tránsito), la cual es alimentada por la policía de tránsito frente a los siniestros ocurridos y se tienen datos.

La información que es subida a la base de datos es la latitud, longitud y dirección del siniestro, la fecha de ocurrencia, formulario y código del accidente, la gravedad, clase, causa, placas e información de los actores viales involucrados.

Esta herramienta permitió cargar y visualizar 586.041 accidentes viales georreferenciados ocurridos entre el año 2007 y 2018 en Bogotá.

El usuario puede filtrar la información por el año, mes, día y hora, mostrando la cantidad de siniestros, georreferenciados en un mapa y generando gráficos estadísticos, en donde se evidencia la cantidad de siniestros hora a hora, por gravedad, clase, número de muertos por días, actor vial muerto y causas de siniestros por días, con el fin de que los expertos puedan analizar las estacionalidades según el tipo de siniestro y los actores involucrados.

Esta sección permite un análisis específico de los siniestros por corredor, con lo cual se puede priorizar los corredores y los tramos en donde más ocurren siniestros para evidenciar posibles causas y asignar recursos que mitiguen los accidentes.

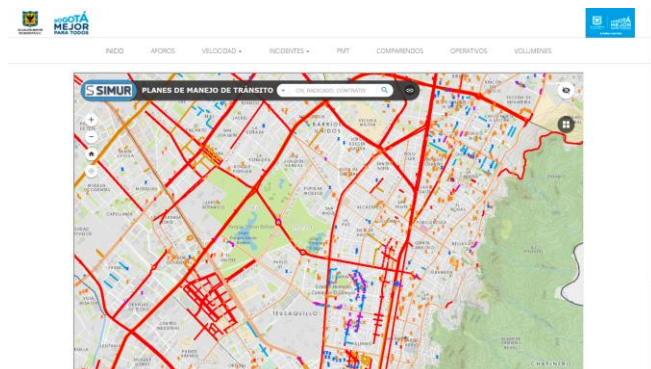


Figura 8. Plataforma SDM, sección de siniestros.  
<http://201.245.176.195:8083/SDMplataforma/siniestralidad.html>

Dentro de esta plataforma encontramos una sección de PMT (Plan de Manejo del Tráfico), en donde se muestran los planes programados de cierres de vía, para así gestionar y

disponer de planes alternativos que permitan una mejor circulación y movilidad, teniendo en cuenta eventos, construcciones, etc que generan cierres en las vías temporalmente.

Esta es una herramienta dispuesta por el Simur (Sistema de Información de Movilidad Urbana), fue desarrollada en ArcGis y embebida dentro de la plataforma, con el fin de integrar toda la información pertinente para la gestión y producción de indicadores, permitiéndole a los expertos identificar si las afectaciones en la vía están autorizadas, y en caso de no serlo proceder con los operativos necesarios para normalizar la circulación.



**Figura 9. Plataforma SDM, sección de PMT.**  
<http://201.245.176.195:8083/SDMplataforma/pmt.html>

Finalmente, la última sección de la plataforma muestra los datos de los comparendos impuestos, donde estos datos constan de latitud, longitud, localidad, dirección, fecha, código y descripción de la infracción. Esta información se carga todo el tiempo por medio de la aplicación interna de la entidad SIMUR (Sistema de Información de Movilidad Urbana) guardando los datos de los comparendos impuestos desde el 2012.

Acá se visualiza la cantidad de comparendos impuestos y los tipos de comparendos georreferenciados a partir de la interacción del usuario con la plataforma, mediante la selección temporal (año, mes, día y hora) de los datos, así como el filtrado espacial (localidad) y la elección de atributos (tipo de comparendo).

Las gráficas que se presentan, muestran la cantidad de comparendos y tipos, por localidad, lo cual presenta una información muy importante con la cual se puede mirar la cantidad de comparendos por tipo de infracción, para así evidenciar cual es la norma que se incumple más para posteriormente hacer campañas ciudadanas entorno a la infracción más recurrente. Así mismo, se puede identificar la localidad y ubicación en donde se generan mayor cantidad de comparendos, con lo cual se identifican estas zonas para así generar mayor presencia de las autoridades con el fin de disminuir el incumplimiento de la norma.

A partir del histórico de los datos, se puede evaluar los días de la semana en que se presentan mayor cantidad de comparendos y así analizar si existe un patrón temporal o un día en específico en donde las personas no cumplen con las normas y evaluar la gestión de las autoridades.



**Figura 10. Plataforma SDM, sección de comparendos.**  
<http://201.245.176.195:8083/SDMplataforma/comparendos.html>

## V. CONCLUSIONES Y TRABAJO FUTURO

La metodología implementada permitió el desarrollo de múltiples tableros, independientemente de los datos o el área de movilidad al cual va dirigido el uso del tablero, ya que todos poseen la misma estructura espacio temporal de los datos, y a partir de su diseño e implementación permitió generar análisis y conclusiones frente al comportamiento de la movilidad reflejado en los datos.

La herramienta web cumplió con el objetivo de implementar el modelo de analítica visual y visualización de gran cantidad de datos heterogéneos, permitiendo al usuario realizar consultas complejas y detalladas mediante una expresividad visual, a través del diseño de la visualización y el uso de controles intuitivos y fáciles de utilizar.

El sistema web es versátil y liviano, por tanto, la visualización y análisis de la información se puede realizar desde un computador hasta en un dispositivo móvil, permitiéndole a los expertos tomar decisiones en campo.

La plataforma permite la integración de datos de distintas áreas de movilidad con lo cual permitió a los expertos consultar múltiples fuentes de datos desde un solo lugar para así proceder a la creación de indicadores y generación de informes frente al comportamiento del transporte en la ciudad.

Como trabajo a futuro se propone aumentar las funcionalidades gráficas para la representación de los resultados a parte del diagrama de barras. Así mismo se desea mostrar la información en tiempo real, mediante la implementación de Tableau Server, con el cual se puede publicar toda la información almacenada y obtenida en “real time”.

## REFERENCIAS

- [1] Plan Maestro de Movilidad – PMM- Indicadores de gestión, Documento ejecutivo, 2017.
- [2] Exploración de datos espacio – temporalis para la analítica visual en sistemas urbanos, Camilo Nemocón. Jose Tiberio Hernandez, 2017.

- [3] Space, time and visual analytics, Gennady Andrienko, Natalia Andrienko, 2010.
- [4] Visualization Data, Ben fry, 2008.
- [5] Interactive Visual tools for spatial multi criteria Decision Making, Gennady Andrienko, Natalia Andrienko, 2002.
- [6] Visual data exploration, Illumining the Path, Cook, 2005.