

## PLATAFORMA DE ANÁLISIS DE DATOS DEL SISTEMA DE TRANSPORTE URBANO DE PASAJEROS

Ponce Sergio Damián; Madrid Francisco; Rubio Debora; Salguero Guadalupe;  
Castañeira Matias

*Universidad Tecnológica Nacional Facultad Regional San Nicolás, Laboratorio GADIB.  
gadib@frsn.utn.edu.ar*

### RESUMEN

En este trabajo se destaca la trascendencia del Sistema Único de Boleto Electrónico (SUBE) en Argentina, señalando su impacto en la movilidad urbana y la calidad de vida. Más allá de simplificar pagos, SUBE ha revolucionado la recopilación y análisis de datos, especialmente en San Nicolás de los Arroyos, Buenos Aires. Este análisis se considera vital para decisiones informadas, mejoras en la experiencia del cliente y la optimización de procesos en el transporte público. El documento ahonda en el modelo de datos, la fuente de datos y resalta la herramienta Qlik BI para generar paneles de control y hojas de trabajo. Los resultados obtenidos, a partir del análisis de 38.1 millones de datos transaccionales, proporcionan información valiosa sobre patrones de viaje, comportamiento del usuario y eficiencia del servicio. Las visualizaciones detalladas, como el gráfico de barras que muestra la fluctuación de pasajeros por mes desde 2018, ofrecen una comprensión intuitiva de eventos significativos, como el impacto de la pandemia en los viajes. Además, el análisis de descuentos en función de la recaudación neta y la evaluación de costos, especialmente en relación con la tarifa plana, subrayan cómo la analítica de datos influye en la toma de decisiones y la planificación del transporte público. Este estudio demuestra cómo SUBE, respaldado por herramientas de análisis de datos se erige como una valiosa fuente de información para la mejora continua de la movilidad urbana.

**Palabras Claves:** Sistema de transporte público, SUBE, analítica de datos, Inteligencia de negocios.

### ABSTRACT

In this work, the significance of the Unique Electronic Ticket System (SUBE) in Argentina is highlighted, emphasizing its impact on urban mobility and quality of life. Beyond simplifying payments, SUBE has revolutionized data collection and analysis, especially in San Nicolás de los Arroyos, Buenos Aires. This analysis is deemed vital for informed decision-making, improvements in customer experience, and the optimization of processes in public transportation. The document delves into the data model, data source, and underscores the Qlik BI tool for generating dashboards and worksheets. The results obtained from analyzing 38.1 million transactional data points provide valuable insights into travel patterns, user behavior, and service efficiency. Detailed visualizations, such as the bar chart depicting passenger fluctuations per month since 2018, offer an intuitive understanding of significant events, such as the impact of the pandemic on travel. Additionally, the analysis of discounts based on net revenue and the evaluation of costs, especially in relation to the flat fare, underscore how data analytics influences decision-making and public transportation planning. This study demonstrates how SUBE, supported by data analytics tools, emerges as a valuable source of information for the continuous improvement of urban mobility.

**Keywords:** Public transportation system, SUBE, data analytics, business intelligence.

## 1. INTRODUCCIÓN

En el mundo actual, el transporte público desempeña un papel crucial en la movilidad urbana y la calidad de vida de millones de personas. En Argentina, el Sistema Único de Boleto Electrónico, conocido como "SUBE," ha revolucionado la forma en que los ciudadanos acceden y utilizan el transporte público. Esta tarjeta electrónica no solo ha simplificado el proceso de pago, sino que también ha abierto un mundo de posibilidades en términos de recopilación y análisis de datos. Como se destaca en Big Data y las implicaciones en la cuarta revolución industrial - Retos, oportunidades y tendencias futuras (G. González, et al) la llamada cuarta revolución industrial presenta oportunidades de congeniar el hardware (en nuestro caso la tarjeta física) de una forma integral que posibilite la recolección y análisis profundo de datos.

El Sistema SUBE, implementado en varias ciudades del país, se ha convertido en un elemento integral de la vida cotidiana de los argentinos, ofreciendo comodidad, eficiencia y ahorro de tiempo en sus desplazamientos diarios. Sin embargo, su importancia trasciende lo meramente operativo. Detrás de cada tarjeta SUBE, se encuentra una verdadera riqueza de información que abarca los patrones de viaje, las preferencias de transporte y los hábitos de movilidad de la población.

En este trabajo de investigación, exploraremos en profundidad la importancia del Sistema SUBE como una herramienta clave en la gestión del transporte público, centrándonos especialmente en la explotación de sus bases de datos utilizando la plataforma Qlik, con un enfoque en la ciudad de San Nicolás de los Arroyos, provincia de Buenos Aires. La recopilación y análisis de datos procedentes de miles de usuarios que utilizan el Sistema SUBE diariamente proporciona información valiosa para la toma de decisiones, la planificación urbana y la mejora de la calidad de los servicios de transporte en esta región específica.

A lo largo de estas páginas, examinaremos cómo la explotación de los datos del Sistema SUBE con la ayuda de una herramienta de BI puede contribuir a la optimización de las redes de transporte público en San Nicolás de los Arroyos. También analizaremos cómo esta información puede desempeñar un papel crucial en la promoción de prácticas de movilidad sostenible y la satisfacción de las necesidades cambiantes de los pasajeros en esta dinámica área urbana.

## 2. DATA ANALYTICS

La era de la información ha transformado radicalmente la forma en que las organizaciones toman decisiones, y no es ajeno a ello el ámbito del transporte público (Sherman, R., 2015). La recopilación y análisis de datos se han vuelto fundamentales para comprender y mejorar los servicios de transporte, y la utilización de herramientas de Business Intelligence (BI) como Qlik ha sido clave en este proceso. En esta sección, exploraremos la importancia y los beneficios de la analítica de datos en el contexto del Sistema SUBE, centrándonos en el enfoque de Business Intelligence y la herramienta Qlik.

La importancia del Data Analytics (Análisis de Datos) en las bases de datos transaccionales es fundamental en el entorno dinámico moderno. Las bases de datos transaccionales son aquellas que registran las operaciones comerciales diarias de una organización, como ventas, compras, registros de clientes y transacciones financieras. Para nuestro objeto de estudio, la base de datos se conforma con los boletos de cada viaje que los pasajeros utilizan. Podemos reconocer algunas de las razones por las cuales el Data Analytics es crucial en este contexto:

**Toma de Decisiones Informadas:** El análisis de datos en bases transaccionales permite a las organizaciones tomar decisiones más informadas y estratégicas. Al explorar y comprender los patrones y tendencias en los datos transaccionales, las organizaciones pueden identificar oportunidades de mejora y optimización.

**Mejora de la Experiencia del Cliente:** El análisis de datos en las bases transaccionales permite a las organizaciones comprender mejor los hábitos y necesidades de las personas. Esto conduce a la personalización de productos y servicios, lo que a su vez mejora la satisfacción del cliente y la retención.

**Optimización de Procesos Internos:** El análisis de datos puede ayudar a identificar ineficiencias en los procesos comerciales. Al analizar los datos transaccionales, las organizaciones pueden encontrar formas de optimizar la gestión de recursos y otros aspectos críticos de su operación.

**Predicción de Tendencias y Demandas:** Es posible predecir tendencias futuras y cambios en la demanda. Esto permite a las organizaciones planificar y ajustar sus operaciones de manera proactiva.

**Cumplimiento Normativo y de Contratos:** Existen exigencias legales y de contrato de servicios que son posibles de verificar con la explotación de los datos almacenados en cada transacción.

En resumen, el Data Analytics desempeña un papel esencial en la optimización de las operaciones comerciales, la mejora de la toma de decisiones y la identificación de estados de calidad de servicio. En las bases de datos transaccionales, el análisis de datos se convierte en una herramienta poderosa para impulsar el crecimiento y la eficiencia de una organización.

## 2.1 Importancia de la Analítica de Datos en el Transporte Público

El transporte público desempeña un papel esencial en la vida urbana, y su eficiencia y calidad afectan directamente la movilidad de los ciudadanos. La analítica de datos se ha convertido en un pilar fundamental para comprender las complejas interacciones entre los usuarios y los sistemas de transporte. En este sentido, el Sistema SUBE recopila una gran cantidad de información valiosa sobre los patrones de viaje, la demanda de servicios y el comportamiento del usuario. La capacidad de analizar estos datos proporciona información que puede utilizarse para optimizar rutas, horarios y recursos, así como para tomar decisiones estratégicas en la planificación del transporte público.

Qlik, una herramienta líder en BI, se ha utilizado con éxito en la explotación de datos del Sistema SUBE. Su capacidad para integrar datos de diversas fuentes, crear visualizaciones interactivas y ofrecer análisis avanzados ha permitido a las autoridades de transporte y planificadores urbanos en San Nicolás de los Arroyos, provincia de Buenos Aires, aprovechar al máximo la información recopilada por el sistema. Esto ha llevado a mejoras notables en la eficiencia y calidad de los servicios de transporte en la región.

En las secciones siguientes de este trabajo, profundizaremos en cómo la analítica de datos y la herramienta Qlik específicamente han influido en la toma de decisiones y la planificación del transporte público en San Nicolás de los Arroyos, demostrando su impacto positivo en la mejora de la movilidad urbana. Además, examinaremos los desafíos y consideraciones éticas relacionados con la explotación de datos en este contexto.

## 3. FUENTE DE DATOS

Tanto las empresas prestadoras de servicio de transporte público como los municipios que poseen el sistema SUBE operativo, reciben de forma mensual un resumen en formato excel con los datos transaccionales. Estos datos se encuentran codificados de forma no muy clara pero es posible ir detectando el significado de sus campos realizando un análisis minucioso. Los datos de estudio y trabajo son los específicos para el Municipio de San Nicolás de los Arroyos en el periodo comprendido entre enero de 2018 a mes agosto de 2023. Los datos fueron disponibilizados a través de un FTP especialmente creado por la agencia Nación Servicios, encargada de la gestión de la administración del sistema SUBE a nivel nacional.

### 3.1 Modelo de datos

Un modelo de datos es una representación estructurada y conceptual de cómo se organiza y almacena la información en una base de datos. Este modelo define cómo se crean, almacenan, relacionan y acceden a los datos dentro de un sistema de gestión de bases de datos.

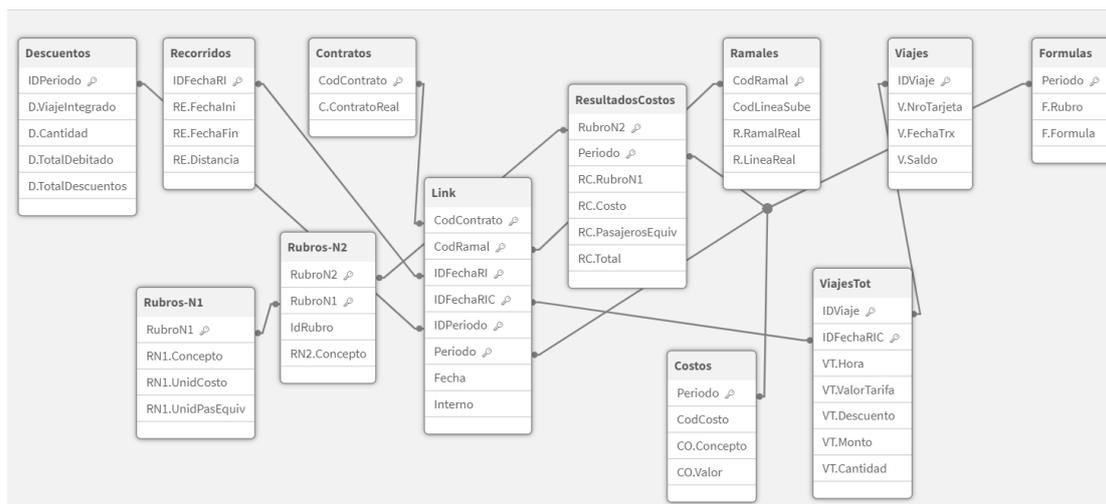


Figura 1 Modelo de datos utilizado.

En nuestro caso se trabajó con 5 archivos proporcionados por el sistema SUBE en formato excel los cuales poseían

- Datos transaccionales de cada boleto
- Integración de datos de recorridos por unidad
- Integración de datos de descuentos
- Integración de datos de cobros
- Integración de datos de servicio

#### 4. HERRAMIENTA DESARROLLADA

Qlik BI es una herramienta de Business Intelligence que simplifica la visualización de datos y permite analizar información de manera interactiva. Carga datos de múltiples fuentes, ofrece seguridad y escalabilidad, y se integra con otras aplicaciones empresariales. Su característica única es la asociación dinámica de datos, lo que facilita el descubrimiento de información relevante. Es una opción valiosa para la toma de decisiones basada en datos y la generación de informes en empresas de todos los tamaños.



##### 4.1 Panel de control y hojas de trabajo

Los paneles de control desempeñan un papel fundamental en una solución de Business Intelligence (BI) debido a su capacidad para ofrecer una vista rápida y visual de los datos clave de una organización. La explotación de datos desarrollada posee un panel de control principal donde se agrupan informaciones de distintas naturalezas como cantidad de viajes, kilómetros recorridos y ranking de líneas de transporte según su uso.

Las distintas hojas de trabajo ahondan en detalle respecto a ámbitos concretos del servicio de transporte público. Fueron desarrolladas 7 pantallas de información específica para:

- Viajes
- Líneas
- Recorrido de internos
- Tarjetas
- Descuentos
- Tarifas
- Costos

Aparte de ello, se generaron hojas de trabajo para realizar escenarios de simulaciones donde analizar el impacto del cambio de tarifa porcentual y su posible impacto en la cantidad de pasajeros.

▼ Hojas públicas (12)



Figura 2 Catálogos de hojas de trabajo.

Los datos recolectados suman 38,1 millones de datos transaccionales (boletos cortados) los cuales corresponden al periodo enero 2018 agosto 2023.



Figura 3 Encabezado de información del dashboard.

#### 4.2 Resultados obtenidos

A lo largo de la navegación de los paneles desarrollados, es posible encontrar información gráfica y de detalle con excelente interpretación incluso para usuarios no relacionados con el negocio del transporte público. Ejemplo de ello lo vemos en la Figura 3 donde se muestran en gráfico de barras la cantidad de pasajeros por mes desde 2018 a la fecha. Podemos observar de forma muy simple el inicio de la pandemia e imaginar su impacto en los valores de recaudación monetaria. En las hojas de detalle de *Viajes* es posible profundizar en esa suposición y verificar una reducción de casi 93%.

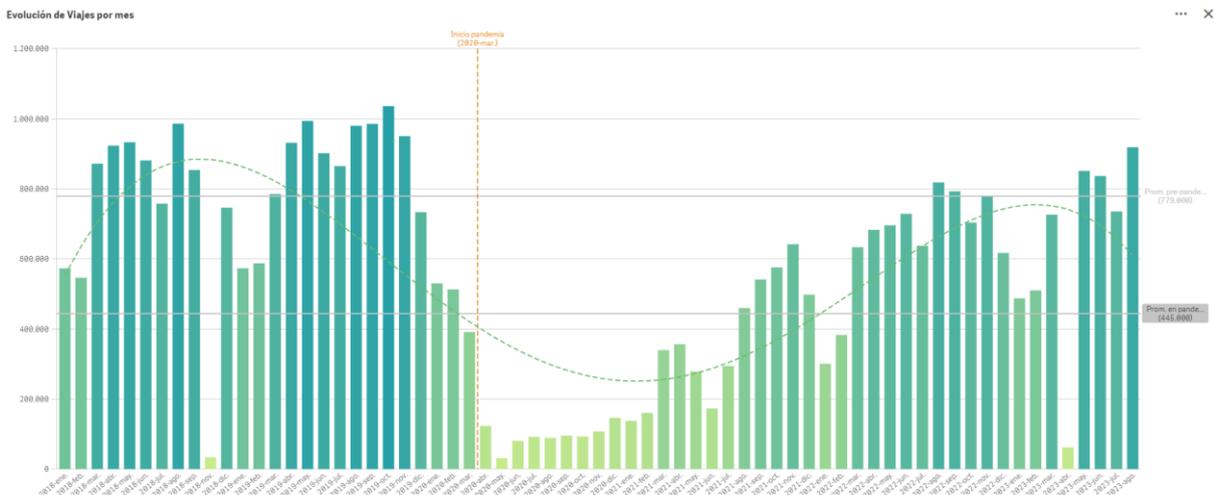


Figura 4 Catálogos de hojas de trabajo.

Es posible encontrar información muy relevante respecto a niveles de *Descuentos* respecto a *Recaudación Neta*. Esta información se hace visible en dimensión a cada línea de transporte en la Figura 4, por la cual podremos interpretar

**Recaudación y descuentos por Línea**

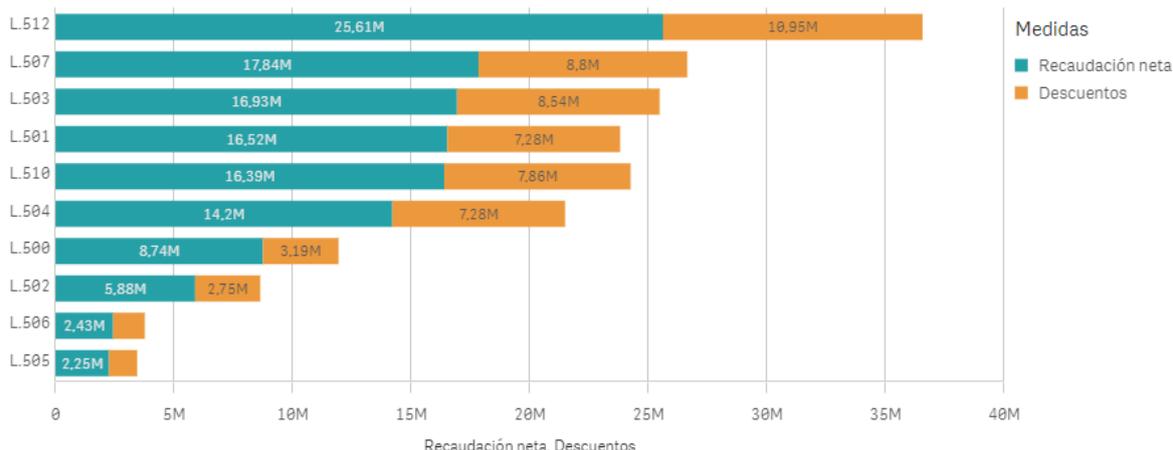


Figura 5 Detalle de ingresos por línea de colectivo.

Una de las secciones más importantes de la captura, análisis y presentación de datos que hemos trabajado está relacionada con los costos del servicio y con su consecuente impacto en la Tarifa Plana. En San Nicolás se utiliza una fórmula polinómica que contempla los costos de los distintos rubros asociados a la prestación del servicio público de transporte, a saber:

<ul style="list-style-type: none"> <li>➊ <b>1. Costos vinculados al kilómetro recorrido</b> <ul style="list-style-type: none"> <li>1.1. Combustible</li> <li>1.2. Lubricantes, filtros y aceites</li> <li>1.3. Desgaste de neumáticos: Neumáticos radiales</li> <li>1.4. Mantenimiento y reparaciones</li> </ul> </li> <li>➋ <b>2. Costos periódicos vinculados a la unidad de transporte</b> <ul style="list-style-type: none"> <li>2.1. Personal de conducción</li> <li>2.2. Amortización del material rodante</li> <li>2.3. Patente</li> <li>2.4. Seguros y previsión de siniestros</li> <li>2.5. Equipamiento del sistema de pago</li> <li>2.6. Sistema de localización automática de vehículos</li> </ul> </li> <li>➌ <b>3. Costos periódicos vinculados a la estructura de la empresa</b> <ul style="list-style-type: none"> <li>3.1. Remuneraciones y cargas sociales del personal administrativo, de playa y de apoyo</li> <li>3.2. Gastos generales</li> <li>3.4. Amortización edificios</li> <li>3.5. Amortización de otros bienes de uso</li> <li>3.6. Juicios no cubiertos por el seguro</li> </ul> </li> <li>➍ <b>4. Costos financieros y de oportunidad</b> <ul style="list-style-type: none"> <li>4.1. Interés sobre el capital invertido</li> <li>4.2. Compensación por el riesgo asumido</li> </ul> </li> </ul>
--

Tabla 1 Rubros y subrubros de costo.

Se creó una interfaz de data entry para la carga manual de cada variable asociada a la ecuación polinómica con 33 variables de cambio mensual y 35 variables estacionarias. El resultado de la aplicación de esta ecuación polinómica presenta a los funcionarios una referencia monetaria del valor de tarifa plana. Es importante destacar esta útil herramienta de análisis ya que su impacto es directo en las negociaciones de costo de boleto con las empresas prestadoras de servicio y su posterior valor de boleto a cada pasajero.

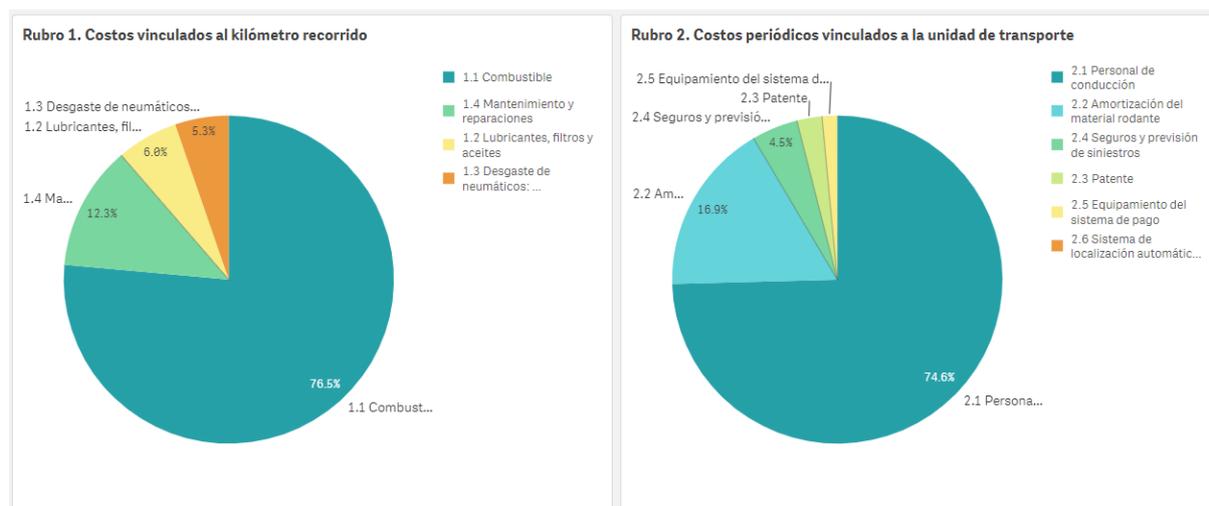


Figura 6 Análisis de costos.

## 5. CONCLUSIONES.

Podemos destacar varias conclusiones a las que hemos llegado con el desarrollo y uso de la herramienta propuesta:

1. **Importancia Estratégica de SUBE:** El Sistema Único de Boleto Electrónico (SUBE) en Argentina no solo simplifica el proceso de pago en el transporte público, sino que se erige como una fuente invaluable de datos. Este estudio destaca su trascendencia en la gestión y mejora continua de la movilidad urbana.

2. **Impacto de la Cuarta Revolución Industrial:** La integración de hardware y software, en este caso, la tarjeta física SUBE y las herramientas de Business Intelligence (BI) como Qlik, permite una recopilación y análisis profundo de datos. Esto responde a las oportunidades presentadas por la llamada cuarta revolución industrial.

3. **Análisis Estratégico con Qlik BI:** La herramienta Qlik BI demuestra ser esencial para explotar la información de SUBE. Sus capacidades para integrar datos, crear visualizaciones interactivas y realizar análisis avanzados han impulsado mejoras notables en la eficiencia y calidad de los servicios de transporte.

4. **Información Clave para la Toma de Decisiones:** Los datos recopilados, que abarcan patrones de viaje, preferencias de transporte y hábitos de movilidad, proporcionan información clave para la toma de decisiones informadas. Esto abarca desde la optimización de rutas y horarios hasta la posible generación de nuevos servicios y eliminación de otros.

5. **Estrategias de Tarifación y Costos:** La aplicación de ecuaciones polinómicas para determinar tarifas planas muestra cómo SUBE y las autoridades pueden utilizar análisis detallados de costos para negociar tarifas con empresas prestadoras de servicios, impactando directamente en los pasajeros.

Existen de igual forma consideraciones éticas y desafíos técnicos importantes. A medida que la analítica de datos se convierte en una herramienta crucial, surgen desafíos éticos relacionados con la privacidad y la seguridad de la información. Es esencial abordar estos aspectos para garantizar un uso responsable de los datos.

En resumen, este trabajo destaca la transformación de SUBE de una tarjeta de transporte a una herramienta estratégica respaldada por análisis de datos avanzados, subrayando su papel en la evolución positiva de la movilidad urbana en San Nicolás de los Arroyos.

#### 4. REFERENCIAS.

Araque González, Gómez Vásquez, M., Vélez Uribe, J. P., & Suárez Hernández, A. H. (2021). *Big Data y las implicaciones en la cuarta revolución industrial - Retos, oportunidades y tendencias futuras*. *Revista venezolana de gerencia*, 26(93), 33–47. <https://doi.org/10.52080/rvg93.04>

Sherman, R. (2015). *Business Intelligence Guidebook*. Amsterdam: Morgan Kaufmann. ISBN: 978-0-12-411461-6

Boquet. (2021). *Contributions to Intelligent Transportation Systems. Big data analytics for reliable and valuable data*.