

# Detrás de los logos: un análisis de la industria automotriz con *Neo4j*

Federico Mazzoni

*Facultad de Ingeniería, Universidad de la República*

Montevideo, Uruguay

luciano.federico.mazzoni@fing.edu.uy

Pablo Hernández

*Facultad de Ingeniería, Universidad de la República*

Montevideo, Uruguay

pablo.hernandez@fing.edu.uy

## Resumen

Este documento presenta el proyecto final realizado en el marco del curso de Bases de Datos No Relacionales de la Facultad de Ingeniería, edición 2023. El objetivo del proyecto es analizar y modelar las relaciones entre los actores de la industria automotriz y sus productos. Con este propósito, se recopila un conjunto de datos sobre vehículos, fabricantes y relaciones entre éstos, y se modela e implementa una base de datos de grafos *Neo4j*, donde se almacenará y visualizará esta información para luego analizarla.

## I. INTRODUCCIÓN

La industria automotriz es un sector complejo, que abarca una amplia variedad de actores que pueden relacionarse entre sí. Empresas matrices, subsidiarias y marcas son algunos de los actores que conforman la red de relaciones en esta industria.

Las empresas matrices desempeñan un papel clave en la industria automotriz. Estas compañías, a menudo de gran envergadura, tienen la propiedad y el control sobre varias marcas y subsidiarias. Establecen la visión estratégica, toman decisiones importantes y definen las políticas comerciales que afectan a todo su grupo de empresas. Como ejemplo de empresa matriz, se puede nombrar a General Motors, empresa norteamericana propietaria de marcas como Chevrolet, Buick, Cadillac, entre otras [1], o al recientemente formado Grupo Stellantis<sup>1</sup>, propietario de Fiat, Ram, Chrysler, Maserati, Peugeot, Alfa Romeo, por nombrar algunas.

Las subsidiarias, por su parte, son empresas filiales controladas por una empresa matriz. Pueden tener una marca propia y operar de manera relativamente independiente, aunque siguen estando bajo el control y la dirección de la empresa matriz. Estas subsidiarias pueden tener roles específicos en términos de producción, investigación y desarrollo, o comercialización de productos. Son las subsidiarias quienes fabrican vehículos que comercializan bajo determinada marca.

Una marca existe con el propósito de asociar una identidad a un conjunto de características en común que tienen los vehículos producidos y comercializados bajo esa marca. Las marcas representan una combinación de historia, reputación, calidad, diseño, tecnología, etcétera. Si se habla de Toyota, se puede hacer alusión a la durabilidad y confiabilidad de los vehículos que produce [2], así como hablar de BMW es hablar de la combinación de lujo y deportividad que ofrecen sus modelos [3]. Otros ejemplos de marcas icónicas incluyen a Ford, Chevrolet, Mercedes-Benz, Mitsubishi, Volkswagen, Audi, entre muchas otras.

Las compañías fabricantes de automóviles pueden tener varias marcas bajo su control, donde cada una está dirigida a un segmento específico del mercado, diferenciado por factores como el precio, el tipo de tecnología que se aplica (por ejemplo autos eléctricos o híbridos), etcétera.

Las compañías pueden relacionarse entre sí, independientemente de si son filiales de la misma empresa matriz, o no. Los casos de colaboraciones entre fabricantes son comunes y variados. Las compañías pueden colaborar para producir piezas, plataformas, e incluso vehículos en conjunto. Es por estas razones, que también existen relaciones entre los vehículos.

Como medida de reducción de costos, los vehículos suelen compartir plataformas. Una plataforma es un grupo consolidado de componentes y sistemas compartidos que conforma todos los aspectos no visuales de un vehículo, por ejemplo, el chasis

<sup>1</sup>About Us — Stellantis <https://www.stellantis.com/en/company/about-us>

(con sus propiedades estructurales), el sistema de calefacción, el motor, y demás. [4]

También en el contexto de la reducción de costos, se ha desarrollado una práctica conocida como *Rebadging*, *Rebranding* o *Badge-Engineering*, que va más allá de la simple utilización de plataformas compartidas. El *Badge-Engineering* (nombrado así irónicamente, haciendo alusión a que la única ingeniería que se realiza es la de cambiar un logo por otro) implica la existencia de autos exactamente iguales, salvo mínimas diferencias de estilo, pero comercializados bajo distinta marca en distintos lugares del mundo. [5]

Existen diversas razones que motivan a realizar *Badge-Engineering*. Como se mencionó anteriormente, al comercializar un vehículo bajo una determinada marca, se establece una asociación entre dicho vehículo y una percepción de calidad o reputación. Así, una manera de poner en práctica el *Badge-Engineering* es tomar un determinado vehículo y comercializarlo bajo varias marcas de la misma empresa, valiéndose de la conexión emocional que tienen los consumidores con una u otra marca, efectivamente logrando que más personas deseen adquirir alguna de las versiones, a pesar de que en el fondo, es en mayor medida el mismo vehículo. Siguiendo esta premisa, un mismo vehículo puede ser comercializado con diferentes nombres en distintas partes del mundo, en función de la imagen de marca que prevalezca en la región. Por ejemplo, varios modelos de automóviles se venden en Australia bajo la marca Holden, ya que esta marca tiene una gran historia y reconocimiento en ese país [6].

Este proyecto se enfoca en las relaciones entre vehículos por medio del *rebranding* y el uso de plataformas en común, así como en las relaciones entre marcas y compañías, y entre diferentes compañías en sí. Con este propósito, se selecciona un conjunto de vehículos que se considera representativo de la industria automotriz, incluyendo autos comercializados tanto a nivel local como a nivel mundial, en diferentes puntos de la historia del automóvil. Se complementa con datos de empresas y las relaciones de los vehículos con ellas. A continuación, se modela una base de datos *Neo4j* con el objetivo de permitir la implementación de consultas que brinden información relevante sobre las relaciones mencionadas.

El artículo se organiza de la siguiente manera. En la sección II se presentan trabajos relacionados al uso de bases de datos de grafos sobre la industria automotriz, así como artículos que abordan la práctica de *rebadging*. La sección III describe el proceso de creación del conjunto de datos utilizado, así como el modelado y la carga de la base de datos. La sección IV detalla las consultas realizadas sobre la base de datos proporcionando una breve descripción, y un análisis de los resultados que retorna cada una. Por último, en la sección V, se presentan las conclusiones del trabajo y se reflexiona sobre posible trabajo a futuro.

## II. TRABAJOS RELACIONADOS

En el artículo “What is an Automotive Platform?” de Motor Trend [7], se explora el concepto de plataforma automotriz y su importancia en la industria.

En relación al *Badge-Engineering*, hay varios recursos disponibles. Los artículos “What is Badge Engineering?” en Autonews Blaster [8] y “Automotive History: Did Badge Engineering Actually Work?” en Curbside Classic [9] ofrecen una visión detallada sobre el tema, analizando ejemplos históricos y evaluando el impacto del *Badge-Engineering* en la industria automotriz. Además, el artículo de Jalopnik [10] presenta una lista de los diez mejores ejemplos de *Badge-Engineering*.

En cuanto a trabajos académicos sobre *Badge-Engineering*, se encontró una escasez de investigaciones en este ámbito. Sin embargo, existen trabajos sobre el uso de grafos en la industria automotriz, como es el caso de [11], donde se exploran las colaboraciones entre empresas a través de un grafo, incluyendo el desarrollo de piezas y *Joint-Ventures*, entre otros tipos de colaboración.

Por otro lado, el artículo “Graphs in Automotive and Manufacturing” en el blog de Neo4j [12], detalla cómo una base de datos de grafos puede representar información variada de la industria automotriz. El artículo presenta cuatro casos de uso fundamentales en esta industria: Gestión de la cadena de suministro, Análisis de garantía, Customer 360 y Grafo de conocimiento. Además, se proporcionan cinco casos de estudio de distintas entidades que han aplicado los grafos en sus operaciones: Volvo Cars, U.S Army, NASA, Schleich y Lockheed Martin.

## III. PARTE CENTRAL

### III-A. Objetivos

Dado que la industria automotriz es un campo muy amplio, la investigación estará dirigida a cuatro puntos específicos. Investigar estos cuatro puntos presentaría una panorama general de la industria, dando lugar también al análisis más específico de la colaboración entre empresas. Los puntos en cuestión son:

- Comparar los lugares de comercialización de los vehículos con los lugares de fundación de sus compañías.
- Establecer una noción de distancia entre vehículos.
- Conocer la compañía más alta en el árbol de relaciones de propiedad para un vehículo.
- Determinar qué vehículos son un *rebadge* de otro vehículo.

### III-B. Elección del tipo de base de datos y modelo

Se decide utilizar una base de datos de grafos debido a su capacidad intuitiva para representar los elementos del universo automotriz y las conexiones entre ellos. Además, el lenguaje de consultas Cypher facilita la obtención de datos basados en estas conexiones, lo cual resulta fundamental para la estructuración del modelo.

Los conceptos seleccionados para representar la industria son:

- *Car* (automóvil)
- *Platform* (plataforma)
- *Brand* (marca)
- *Company* (compañía)
- *Country* (país)

en conjunto con las relaciones:

- *Rebrand* (reflexiva para *Car*)
- *Makes* (entre *Brand* y *Car*)
- *Based on* (entre *Car* y *Platform*)
- *Develops* (entre *Company* y *Platform*)
- *Owns* (reflexiva para *Company* y también entre *Company* y *Brand*)
- *Assembled in* (entre *Car* y *Country*)
- *Headquarters* (entre *Company* y *Country*)
- *Was part of* (reflexiva para *Company*)

Con estas entidades y relaciones es posible representar dónde se fabrican los vehículos, el hecho de que algunos vehículos son *rebadge* entre sí, conocer la marca bajo la cual se comercializan, qué compañía es propietaria de otra compañía o de una marca, las plataformas que se desarrollan y utilizan para los vehículos, y a qué país pertenece una compañía. La [Figura 1](#) contiene el diagrama del modelo que se define para representar los conceptos y sus relaciones.

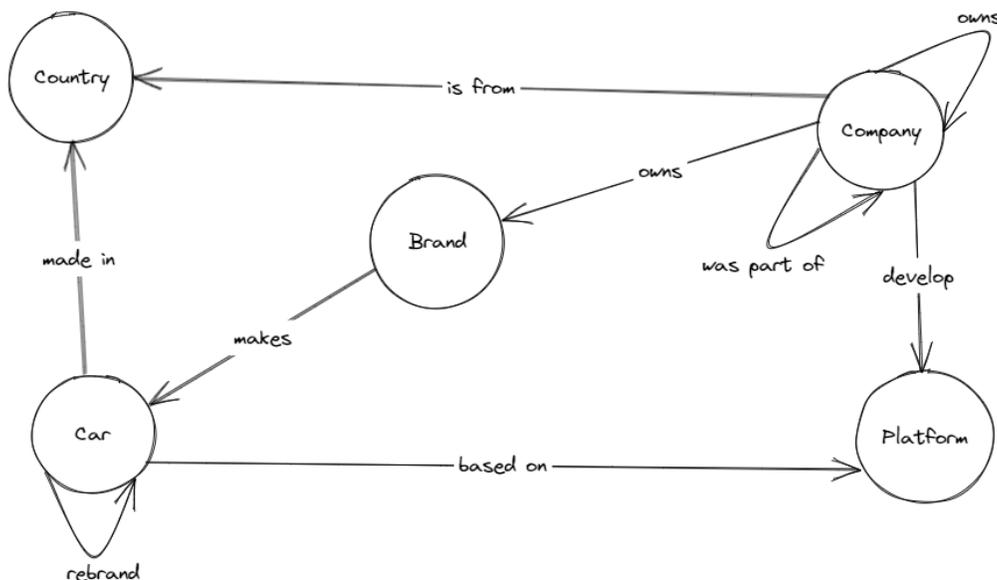


Figura 1: Diagrama del modelo.

### III-C. Obtención de datos

Los datos utilizados incluyen modelos del mercado local, específicamente vehículos nuevos disponibles en el primer semestre de 2023, obtenidos de *autoblog*<sup>2</sup>. Además, se complementa con ejemplos de *Badge-Engineering* extraídos de *Donut Media*<sup>3</sup>, cuya investigación se puede ver en video en [13]. También se tomaron como referencia ejemplos de [14] y de [15]. Para complementar aun más, el equipo se contactó con el propietario de este último canal, conversación que condujo a consolidar aun más la idea de buscar vehículos que compartieran plataforma. Por último, también se incorporaron vehículos basados en conocimientos previos del equipo.

El conjunto de datos resultante se compone de cuatro archivos .csv con información sobre autos, compañías, marcas y plataformas, respectivamente, y consta de 90 vehículos, 34 compañías, 38 marcas, 34 plataformas y 45 países.

Para cada auto se tiene la siguiente información:

- Marca
- Modelo
- Carrocería
- Generación<sup>4</sup> (nombre u ordinal)
- Plataforma utilizada
- Países de fabricación/ensamblaje
- Año de comienzo de producción (de la generación)
- Año de fin de producción (de la generación)
- Vehículos con los que se relaciona por medio de un rebranding

El *rebranding* se da en grupos de autos, dentro de los cuales todos los autos son un *rebrand* del resto de autos de ese grupo; para simplificar la representación, se eligió arbitrariamente un vehículo dentro de cada grupo de *rebranding*, y todas las relaciones (las cuales son bidireccionales) del grupo apuntan a este nodo central. Un ejemplo de esto se puede ver en la [Figura 2](#).



Figura 2: Grupo de *rebadge* del Izuzu Gemini.

Para cada compañía se tiene:

- Nombre de la compañía

<sup>2</sup>Portal dedicado al periodismo de automóviles en Uruguay, especialmente enfocado en noticias y reseñas de automóviles vendidos en el país. <https://www.autoblog.com.uy/>

<sup>3</sup>Empresa de contenido digital automotriz estadounidense <https://donut.media/>

<sup>4</sup><https://www.cargurus.com/Cars/articles/understanding-vehicle-generations>

- Año de fundación
- País de las oficinas centrales
- Compañía dueña (si aplica)
- Empresas de las que formó parte en el pasado

Para cada marca se tiene:

- Nombre de la marca
- Compañía que la controla

Para cada plataforma se tiene:

- Nombre de la plataforma
- Empresas que participaron en su desarrollo

Para cada país se conoce su nombre.

Los datos de cada archivo se obtienen de fuentes variadas como [16], [17], [18], y [19]

### III-D. Creación de la base de datos

Se trabajó con una instancia de base de datos local creada en *Neo4j Desktop*. Este software incluye a *Neo4j Browser*, una interfaz gráfica que permite visualizar los nodos y vínculos de la base de datos además de ofrecer una línea de comandos que se utilizó para cargar la base de datos. *Neo4j Desktop* también permite instalar el *plugin Graph Data Science Library*, utilizado para detectar comunidades y determinar la centralidad de los nodos, lo que se verá en la sección IV.

Con el objetivo de disponibilizar el proyecto, se creó una instancia de la base de datos en *Neo4j Aura* [20], accesible a través de un archivo de *Google Colaboratory*, adjunto con el proyecto.

### III-E. Carga de datos

Para cargar los datos en *Neo4j* se utiliza un script escrito en *Cypher Query Language* (CQL), ingresado en la línea de comandos de *Neo4j Browser*. Este script se divide en 4 comandos *LOAD*, uno para cada archivo .csv. Cada *LOAD* se encarga de recorrer las filas de un archivo .csv en particular, creando nodos en la base de datos con los datos de las columnas como atributos del nodo, y creando las asociaciones que correspondan para las columnas que se traten de asociaciones. En [1] se puede observar el código utilizado para cargar las compañías en la base de datos. El script antes mencionado, al igual que los csv's empleados se pueden encontrar en el repositorio del proyecto [21].

#### Listado 1 Script CQL donde se cargan las *Companies*

---

```
LOAD CSV WITH HEADERS FROM 'file:///companies.csv' AS row
MERGE (company:Company {name: row.name})
WITH company, row.headquarters AS headquarters
MERGE (country:Country {name: headquarters})
MERGE (company)-[:HEADQUARTERS]->(country);
```

---

## IV. EXPERIMENTACIÓN

Esta sección está dedicada al análisis del grafo construido en la parte anterior, realizando consultas en CQL orientadas a exponer casos de interés e intentar inferir información sobre el comportamiento de las compañías a partir de sus relaciones y las de los vehículos que producen. Para estos comportamientos, se plantean hipótesis que intentan explicarlos. Adicionalmente, se ejecutan algoritmos de detección de comunidades y de centralidad sobre el grafo. Los resultados de las consultas que tienen formato tabla se incluyen en el archivo de *Google Colaboratory* mencionado previamente, además de incluir en este documento las partes más relevantes de estos resultados.

### IV-A. Consultas y sus resultados

Previo a las consultas, se introduce el término de “compañía dueña directa” que corresponde a la primera empresa encontrada al recorrer el grafo por medio de relaciones *Owns* partiendo desde un auto. En otras palabras, estas son las empresas subsidiarias de más bajo nivel en el grafo de relaciones de propiedad entre las compañías, que también son las empresas dueñas de las *Brands* o marcas.

*IV-A1. Autos, lugares de ensamblaje y país de origen de la empresa matriz:* Esta consulta está orientada a intentar obtener cierta información sobre el alcance de las empresas matrices en el mundo. Los datos que se tienen disponibles para obtener alguna conclusión son los lugares de ensamblaje de los autos, por lo que es de interés comparar estos con el país de origen de una empresa matriz.

Con esta consulta se intenta abordar el objetivo planteado de comparar los lugares de origen de la compañía que produce un vehículo con el lugar de venta de éstos. Si bien no se tienen datos sobre el lugar de venta de un vehículo, igualmente se puede inferir información similar al respecto conociendo los lugares de ensamblaje.

Car	Company	Parent company origin	Assembly co...	Assembled in
"Honda Civic"	"Honda Motor Company"	"Japan"	13	["Canada", "Pakistan", "Malaysia", "Turke...
"BMW 3-Series"	"BMW Group"	"Germany"	10	["Egypt", "Thailand", "Indonesia", "Chin...
"Mercedes-Benz C Class"	"Mercedes-Benz Group"	"Germany"	10	["India", "Vietnam", "Indonesia", "German...
"Opel Astra"	"Stellantis"	"Netherlands"	10	["Russia", "Germany", "Egypt", "South Afri...
"Hyundai Tucson"	"Hyundai Motor Company"	"South Korea"	9	["Egypt", "Czech Republic", "Taiwan", "So...
"Toyota Corolla"	"Toyota Motor Corporation"	"Japan"	8	["Turkey", "China", "Brazil", "Japan", "Taiw...
"Daewoo Gentra"	"General Motors"	"USA"	7	["Ecuador", "Egypt", "Colombia", "Ukrain...
"Chevrolet Aveo"	"General Motors"	"USA"	7	["Ecuador", "South Korea", "India", "Egyp...

Figura 3: Parte de la tabla de resultados de la consulta 1.

Observando los resultados en la [Figura 3](#), se puede notar que la cantidad de países de origen de las empresas matrices es significativamente menor en comparación con la cantidad de lugares de ensamblaje utilizados por estas empresas para producir sus vehículos. Mientras que la mayoría de las empresas se ubican en Europa, Estados Unidos, Japón y Corea del Sur, los lugares de ensamblaje se encuentran prácticamente en todas las regiones del planeta. Esto refleja las estrategias de producción global adoptadas por las empresas y posiblemente la popularidad de los vehículos. Es interesante destacar que el Honda Civic es el automóvil producido en más lugares, con un total de 13 países. Esto sugiere una demanda significativa de este modelo, lo cual justificaría la distribución de su producción a lo largo de varios continentes.

*IV-A2. Cantidad de marcas distintas que poseen las filiales:* Esta consulta apunta a obtener información sobre la variedad de vehículos y marcas que puede llegar a ofrecer una empresa sin necesidad de fusionarse con otras para poseer en conjunto más marcas.

Company	Brands	Brand count	Cars	Car count
"General Motors"	["Chevrolet", "Pontiac", "Daewoo...]	5	["Chevrolet Onix", "Chevrolet Av...]	12
"Stellantis"	["Peugeot", "Dodge", "Ram", "Je...]	4	["Peugeot 308", "Peugeot 208", ...]	10
"BMW Group"	["BMW", "Mini"]	2	["BMW 3-Series", "BMW X3", "M...]	3
"Toyota Motor Corporation"	["Toyota", "Lexus"]	2	["Toyota Land Cruiser", "Toyota ...]	11
"Hyundai Motor Company"	["Hyundai", "Kia"]	2	["Hyundai HB20", "Hyundai Tucs...]	4

Figura 4: Parte de la tabla de resultados de la consulta 2.

Observando los resultados en la [Figura 4](#), General Motors se posiciona como la empresa dueña directa con la mayor cantidad de marcas bajo su control, seguida por Stellantis. Como se mencionó en la introducción, tener varias marcas permite a estas empresas producir modelos variados o incluso vender un mismo modelo con modificaciones menores bajo cada una de las distintas marcas, con el objetivo de abarcar la mayor cantidad posible de segmentos del mercado.

*IV-A3. Plataformas, autos que las usan y empresas matrices que las desarrollan:* Esta consulta busca dar magnitud a la colaboración entre empresas matrices por medio del desarrollo en conjunto de plataformas. A su vez, muestra una medida cuantitativa de cuánto se reutilizan las plataformas, utilizando la cantidad de vehículos basados en ella.

Platform	Companies	Company Count	Cars	Car Count
"T6"	["Mazda Motor Corporation", "V..."]	3	["Mazda BT-50", "Volkswagen Am..."]	3
"T-Platform"	["General Motors", "Stellantis", "..."]	3	["Pontiac Acadian", "Buick Opel"..."]	12
"GS"	["Stellantis", "Mitsubishi Motors"]	2	["Dodge Journey", "Mitsubishi L..."]	3
"MLB Evo"	["Volkswagen Group"]	1	["Audi Q8", "Lamborghini Urus"]	2
"MC-P"	["Stellantis"]	1	["Fiat Strada", "Ram 700"]	2
"EMP2"	["Stellantis"]	1	["Peugeot 308", "Opel Astra L"]	2

Figura 5: Parte de la tabla de resultados de la consulta 3.

Los resultados de la [Figura 5](#) denotan que la colaboración entre empresas matrices no es una práctica tan común. Dentro del conjunto de datos, solo se identificaron tres plataformas utilizadas por más de una empresa matriz. Sin embargo, resulta interesante destacar la presencia de General Motors en el segundo puesto. A pesar de ser la compañía con mayor cantidad de marcas bajo su control, forma parte de su estrategia comercial el desarrollo de plataformas en colaboración con otras empresas. Por otro lado, el primer puesto está ocupado por tres empresas, lo cual indica un alto grado de reutilización de una plataforma común. Además, al investigar la cantidad de países en los que se ensamblan vehículos que utilizan la plataforma T6, se observa que abarca países de cuatro continentes ([Figura 6](#)).

Platform	Cars	Countries of Assembly	Country count
"T6"	["Mazda BT-50", "Volkswagen Amarok", "Ford Ranger"]	["South Africa", "Thailand", "Nigeria", "Vietnam", "Argentina", "USA", "Cambodia"]	7

Figura 6: Tabla de resultados de la consulta 3.5.

*IV-A4. Rebadge dentro de una empresa matriz:* Esta consulta muestra cuantitativamente la aplicación de la práctica del *rebadge* dentro de una misma empresa matriz. Muestra pares de autos vinculados a través de un *rebadge* y que pertenecen a la misma empresa matriz.

Esta es una de las consultas que abordan el objetivo planteado de conocer los *rebadges* de cada auto.

Car 1	Car 2	Company
"Pontiac G3 Wave"	"Chevrolet Aveo"	"General Motors"
"Daewoo Gentra"	"Chevrolet Aveo"	"General Motors"
"Pontiac T1000"	"Chevrolet Chevette"	"General Motors"
"Pontiac Acadian"	"Chevrolet Chevette"	"General Motors"

Figura 7: Parte de la tabla de resultados de la consulta 4.

Observando los resultados (parte de ellos en la [Figura 7](#)), se puede apreciar que General Motors es la empresa que realiza la mayor cantidad de *rebadging* en sus propios vehículos, lo cual respalda la hipótesis planteada de que controlar múltiples marcas y utilizar los mismos vehículos a través de todas ellas contribuye a aumentar las ventas.

*IV-A5. Rebadge entre diferentes compañías dueñas directas:* Esta consulta muestra cuantitativamente la aplicación de la práctica del *rebadge* entre compañías dueñas directas. Muestra pares de autos vinculados a través de un *rebadge* y que pertenecen a distintas empresas dueñas directas. Al igual que la consulta anterior se relaciona indirectamente con el objetivo planteado sobre *rebadges*.

Car 1	Car 2	Company 1	Company 2
"Chevrolet San Remo"	"Holden Gemini"	"General Motors"	"GM Holden Ltd"
"GMC Chevette"	"Holden Gemini"	"General Motors"	"GM Holden Ltd"
"Pontiac T1000"	"Holden Gemini"	"General Motors"	"GM Holden Ltd"
"Pontiac Acadian"	"Holden Gemini"	"General Motors"	"GM Holden Ltd"
"Chevrolet Chevette"	"Holden Gemini"	"General Motors"	"GM Holden Ltd"

Figura 8: Parte de la tabla de resultados de la consulta 5.

Nuevamente, interesa notar en los resultados de esta consulta (parte de ellos en la [Figura 8](#)) la fuerte presencia de General Motors en su relacionamiento con el resto de empresas de la industria, ubicándose como protagonista en varias relaciones de *rebadge*.

*IV-A6. Rebadge entre empresas matrices:* Esta consulta analiza la colaboración entre empresas matrices por medio del *rebadging* y también aborda de manera implícita el objetivo planteado sobre encontrar los *rebadges* entre autos, mostrando los pares de autos vinculados a través de un *rebadge* pero que tienen distinta empresa matriz.

Car 1	Car 2	Company 1	Company 2
"Vauxhall Chevette"	"Chevrolet San Remo"	"Stellantis"	"General Motors"
"Opel Kadett"	"Chevrolet San Remo"	"Stellantis"	"General Motors"
"Vauxhall Chevette"	"Pontiac T1000"	"Stellantis"	"General Motors"
"Opel Kadett"	"Pontiac T1000"	"Stellantis"	"General Motors"
"Vauxhall Chevette"	"Pontiac Acadian"	"Stellantis"	"General Motors"

Figura 9: Parte de la tabla de resultados de la consulta 6.

Los resultados ([Figura 9](#)) reflejan otra particularidad de la industria automotriz: su dinamismo. Si bien se incluyen varios modelos donde se pueden ver *rebadges* que protagonizan General Motors y Stellantis, varios de estos se dieron cuando empresas

subsidiarias de Stellantis formaban parte de General Motors. De todas maneras, se pueden ver algunas colaboraciones reales entre empresas matrices, aunque en menor medida que para el caso de *rebadges* dentro de una misma empresa matriz.

*IV-A7. Caminos desde un auto a su empresa matriz:* Esta consulta muestra los caminos del grafo que empiezan desde un auto y recorren las relaciones de pertenencia *Owns* hasta llegar a la empresa que termina siendo responsable de la fabricación de ese auto.

Esta consulta responde al objetivo de encontrar el camino desde un auto hasta la compañía más alta en el árbol de relaciones de propiedad.

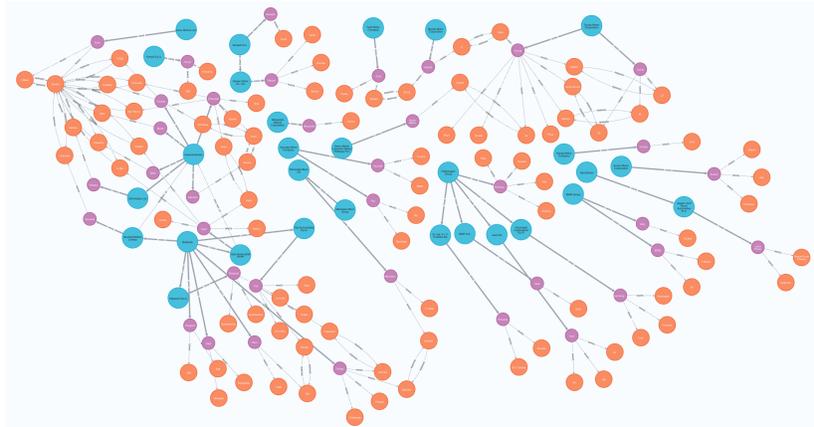


Figura 10: Grafo que representa los caminos de los autos a sus empresas matrices.

Se puede observar en la [Figura 10](#) que la mayoría de los vehículos del grafo tienen una ruta de relaciones a través de la propiedad *Owns* que finaliza en un número limitado de empresas. Stellantis, General Motors y Volkswagen Group son las empresas con mayor centralidad en este sentido. Es natural que estas empresas sean las últimas responsables de la fabricación de un vehículo, ya que son grupos empresariales que controlan un número considerable de marcas.

*IV-A8. Grupos de rebadging en el grafo:* Esta consulta permite visualizar los distintos grupos de *rebadging* que se encuentran presentes en el grafo.

Se observa en la [Figura 2](#) el grupo de *rebadging* más grande de todo el grafo, que tiene al Izuzu Gemini como nodo central. Con esta consulta se aborda el objetivo de encontrar los *rebadges* para cada auto, representando visualmente los conjuntos de autos que forman alguna relación de *rebadge*.

*IV-A9. Centralidad de compañías:* Con el objetivo de analizar aun más la relevancia de las compañías, se ejecuta el algoritmo *Degree Centrality* [22] sobre el grafo. Este algoritmo determina la popularidad de los nodos del grafo, dependiendo de las relaciones entrantes y/o salientes que tenga el nodo. Se ejecutó el algoritmo con orientación *UNDIRECTED* para las relaciones. Es importante destacar que el algoritmo se ejecuta teniendo en cuenta todos los nodos y relaciones del grafo, pero luego se filtra por tipo de nodo para mostrar los puntajes de las compañías.

	Company	score
1	"General Motors"	12.0
2	"Stellantis"	12.0
3	"Volkswagen Group"	10.0
4	"Hyundai Motor Company"	7.0
5	"Fiat Chrysler Automobiles N.V."	7.0
6	"Nissan Motor Co. Ltd."	6.0
7		

Figura 11: Centralidad de las compañías.

Se observa en la [Figura 11](#) que las compañías con mayor centralidad son Stellantis y General Motors, seguidas por Volkswagen Group, resultado que no difiere de lo esperado, puesto que son las compañías de mayor tamaño. Nota: este resultado también se ve afectado por relaciones de propiedad que tuvieron las compañías en el pasado, a través de la relación *Was part of*, por lo que no solamente contempla la realidad actual de las empresas.

*IV-A10. Centralidad de autos:* Se ejecuto el mismo algoritmo de centralidad para detectar la centralidad de los autos en el grafo.

	Car	score
1	"Izuzu Gemini"	25.0
2	"Honda Civic"	14.0
3	"Opel Astra"	13.0
4	"Chevrolet Aveo"	12.0
5	"Mercedes-Benz C Class"	12.0
6	"Ford Ranger"	11.0
7		

Figura 12: Centralidad de los autos.

Se observa en la [Figura 12](#) que el Izuzu Gemini está primero, lo cual tiene sentido, ya este es el modelo con más *Rebadges*. Tampoco sorprende la presencia del Honda Civic en el segundo puesto, visto que es el auto con más lugares de ensamblaje.

*IV-A11. Detección de comunidades:* Se utilizó el algoritmo de detección de comunidades *Strongly Connected Components* [23] para identificar los grupos de *Rebadging*. Este resultado se presentó previamente de manera visual con la consulta *Grupos de rebadging en el grafo*, pero el algoritmo de detección de comunidades permite obtener la lista de autos en cada comunidad (grupo de *rebadging*). Se muestran parte de los resultados en la [Figura 13](#).

Component	Cars
28	["Izuzu Gemini", "Buick Opel", "Izuzu I-Mark", "Holden Gemini", "GMC Chevette", "Opel K180", "Chevrolet Chevette", "Pontiac T1000", "Pontiac Acadian", "Chevrolet San Remo", "Vauxhall Chevette", "Opel Kadett"]
2	["Chevrolet Aveo", "Pontiac G3 Wave", "Daewoo Genra"]
5	["Dodge Journey", "Fiat Freemont"]
7	["Ford Ranger", "Mazda BT-50"]
10	["Toyota Land Cruiser", "Lexus LX"]
12	["Toyota Altezza", "Lexus IS"]
14	["Toyota Celsior", "Lexus LS"]
16	["Ram 700", "Fiat Strada"]
18	["Aston Martin Cygnet", "Toyota iQ"]
22	["Opel Astra", "Chevrolet Astra"]
24	["Toyota Yaris", "Mazda 2"]
26	["Mercedes-Benz Sprinter", "Dodge Sprinter"]
0	["Lamborghini Urus"]
1	["Audi Q8"]
9	["Volkswagen Amarok"]
20	["Peugeot 308"]
21	["Opel Astra L"]
40	["Chevrolet Onix"]
41	["Chevrolet Tracker"]
42	["Fiat Uno Way"]

Figura 13: Grupos de *rebadge*.

## V. CONCLUSIONES Y TRABAJO FUTURO

### V-A. Conclusiones

Las bases de datos no relacionales ofrecen una solución eficiente y versátil para almacenar y consultar datos voluminosos y no estructurados. En particular, las bases de datos de grafos se destacan por su capacidad para representar y modelar relaciones complejas entre entidades, lo que las convierte en una opción ideal para dominios donde las relaciones son fundamentales.

En el contexto de este proyecto, la elección de una base de datos de grafos resultó especialmente apropiada para modelar las complejas relaciones entre los conceptos. Por ejemplo, se pueden establecer relaciones de propiedad entre compañías y marcas, relaciones de producción entre marcas y vehículos, y relaciones entre vehículos a través de prácticas como el *rebadging*.

El análisis del grafo construido permitió obtener conclusiones interesantes sobre el comportamiento de las empresas de la industria automotriz en su búsqueda de estrategias para aumentar las ganancias. Por ejemplo, se observó la distribución global de la producción de vehículos para reducir costos y satisfacer la demanda en diferentes regiones. También se identificaron asociaciones estratégicas entre empresas para producir plataformas en conjunto o aprovechar modelos existentes para comercializarlos bajo su propia marca.

Este proyecto se basó en el uso de *Neo4j* y el lenguaje de consultas *Cypher Query Language*, herramientas que resultaron fundamentales para la creación y análisis de la base de datos de grafos. La estructura de la base de datos permitió almacenar

y relacionar de manera efectiva los diversos conceptos requeridos.

Durante el proceso de creación de la base de datos, se destacó la importancia del fenómeno del *rebadging* en la industria automotriz. Aunque la literatura académica sobre este tema específico es limitada, se exploraron trabajos y análisis relacionados que proporcionaron una comprensión más amplia del fenómeno, lo cual contribuyó a abordar el tema desde una perspectiva amplia y considerar diferentes fuentes de información.

Todos los objetivos inicialmente planteados fueron logrados, con excepción de establecer una noción de distancia entre vehículos, lo cual quedó representado implícitamente en la sección IV para las consultas *Cantidad de marcas distintas que poseen las filiales y Plataformas, autos que las usan y empresas matrices que las desarrollan*, ya que, más allá de no cuantificar explícitamente la distancia, identifican una relación indirecta entre los autos, al compartir ya sea plataformas o marcas, lo cual es una relación con una distancia mayor que la relación más directa del *rebadge*.

En conclusión, este informe demuestra la viabilidad de utilizar *Neo4j* para crear una base de datos sobre la industria automotriz y aplicar distintos métodos para analizarlo e inferir información relevante sobre la misma, siendo estos métodos la realización de consultas en *Cypher*, y la ejecución de algoritmos de centralidad y de detección de comunidades. También destaca la importancia del fenómeno del *rebadging* en la industria automotriz y la necesidad de llevar a cabo investigaciones más rigurosas en este campo.

#### V-B. Trabajo futuro

Una posible mejora para el modelo en el futuro sería incorporar una nueva relación entre *Car* y *Country* para conectar los vehículos con los lugares donde se comercializan, lo que permitiría aprovechar al máximo los conceptos existentes, ampliando sus conexiones. Esta adición podría dar lugar a consultas que revelen información interesante, por ejemplo, al obtener los lugares de ensamblaje y venta de un vehículo, se podría determinar qué tan común es que un automóvil se ensamble en el país donde se vende. Además, en el caso del *rebadging*, se podrían consultar los automóviles que se relacionan entre sí y que a la vez se comercializan en el mismo lugar, lo cual sería una estrategia de comercialización llamativa, ya que en una misma región se estarían ofreciendo dos vehículos idénticos pero con marcas y nombres diferentes. Esto abriría la puerta a nuevas investigaciones y análisis en torno a la influencia psicológica de las marcas en los consumidores y sus decisiones de compra.

Por otro lado, resultaría interesante considerar de alguna manera los años de producción de los vehículos, así como ampliar el conjunto de datos con información histórica sobre la propiedad y fusiones de empresas. Aunque se intentó representar estas situaciones con la relación *Was part of*, se podría profundizar en esta temática para reflejar de manera más precisa tanto la realidad actual de la industria automotriz como su evolución a lo largo de los años.

#### REFERENCIAS

- [1] List of subsidiaries of General Motors Company <https://www.sec.gov/Archives/edgar/data/1467858/000119312510192195/dex211.htm>
- [2] Who Makes the Most Reliable New Cars? <https://www.consumerreports.org/car-reliability-owner-satisfaction/who-makes-the-most-reliable-cars-a7824554938>
- [3] Mastering The Road: The Iconic Journey Of BMW <https://thebrandhopper.com/2023/05/17/mastering-the-road-the-iconic-journey-of-bmw/>
- [4] What is an Automotive Platform? A Crash Course <https://www.magna.com/stories/inside-automotive/platform/automotive-platform>
- [5] Same Car, Different Brand, Hugely Higher Price: Why Pay An Extra \$30,000 For Fake Prestige? <https://www.forbes.com/sites/eamonnfingleton/2013/07/04/same-car-different-brand-hugely-higher-price-why-pay-an-extra-30000-for-fake-prestige/?sh=4e7ac1117407>
- [6] What is badge engineering? <https://www.carsales.com.au/editorial/details/what-is-badge-engineering-27518/>
- [7] Motor Trend: "What is an Automotive Platform?" <https://www.motortrend.com/features/what-is-an-automotive-platform/>
- [8] Autonews Blaster: "What is Badge Engineering?" <https://autonewsblaster.com/2019/11/what-is-badge-engineering-enthusiast-101.html>
- [9] Curbside Classic: "Automotive History: Did Badge Engineering Actually Work?" <https://www.curbsideclassic.com/automotive-histories/automotive-history-did-badge-engineering-actually-work/>
- [10] Jalopnik: "The Ten Best Examples of Badge Engineering" <https://jalopnik.com/the-ten-best-examples-of-badge-engineering-1536869332>
- [11] Neo4j Car Manufacturers <https://neo4j.com/graphgists/car-manufacturers-2013/>
- [12] Neo4j Blog: "Graphs in Automotive and Manufacturing" <https://neo4j.com/blog/graphs-in-automotive-and-manufacturing/>
- [13] Donut Media: "Who Really Made Your Car: The Truth About Badge Engineering" [https://www.youtube.com/watch?v=3MTf1QZmyxEt=2sab\\_channel=Donut](https://www.youtube.com/watch?v=3MTf1QZmyxEt=2sab_channel=Donut)
- [14] Wikipedia: "List of badge-engineered vehicles" [https://en.wikipedia.org/wiki/List\\_of\\_badge\\_engineered\\_vehicles](https://en.wikipedia.org/wiki/List_of_badge_engineered_vehicles)
- [15] Ed's Auto Reviews: "Ep. 30 Seeing Double: The History of Badge Engineering" [https://youtu.be/H69jd12q\\_0U](https://youtu.be/H69jd12q_0U)
- [16] Autodata <https://www.auto-data.net/es/>
- [17] Wikipedia <https://en.wikipedia.org/>
- [18] Consumer Reports: "Who Owns Which Car Brands?" <https://www.consumerreports.org/cars/who-owns-which-car-brands-a5925868661/>
- [19] Back4App: Car Models by Manufacturer, Category and Year <https://www.back4app.com/database/back4app/car-make-model-dataset>
- [20] Neo4j Aura <https://console.neo4j.io/>
- [21] Repositorio del proyecto en Gitlab <https://gitlab.fing.edu.uy/pablo.hernandez/bases-de-datos-no-relacionales/>
- [22] Degree Centrality <https://neo4j.com/docs/graph-data-science/current/algorithms/degree-centrality/>
- [23] Strongly Connected Components <https://neo4j.com/docs/graph-data-science/current/algorithms/strongly-connected-components/>