**Gestión de Datos – TP1C2022 – Telemetría Carreras**

**Entrega III – Modelo de BI y carga de datos.**

**Primera Reentrega**

Grupo: **MARCO\_AND\_FRIENDS.**

Curso: **K3012.**

Integrantes: **Cabaleiro, Ulises 1777361.**

**Piloni, Ignacio 1764445.**

**Rodriguez, Mora 1763908.**

**Piatti Castro, Marco Augusto 1756450.**

Indice

[Modelo de Datos Transaccional propuesto: 3](#_Toc108029024)

[Estrategia de Migración del modelo Transaccional: 4](#_Toc108029025)

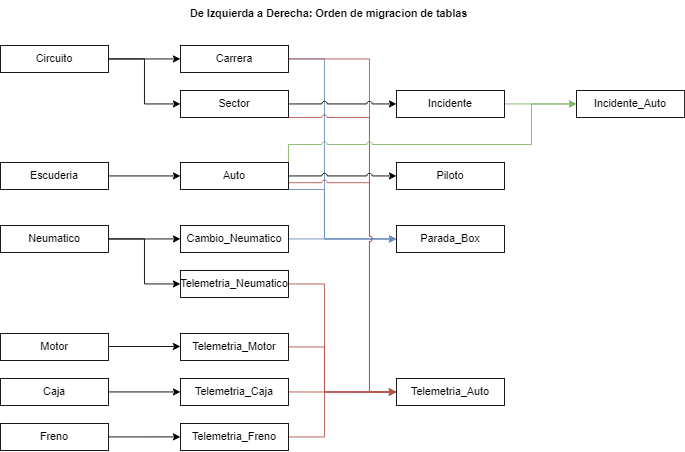
[Modelo de datos BI: 5](#_Toc108029026)

[Estrategia de migración del modelo BI: 5](#_Toc108029027)

[Algunas Aclaraciones Puntuales: 6](#_Toc108029028)

# Modelo de Datos Transaccional propuesto:

# Estrategia de Migración del modelo Transaccional:

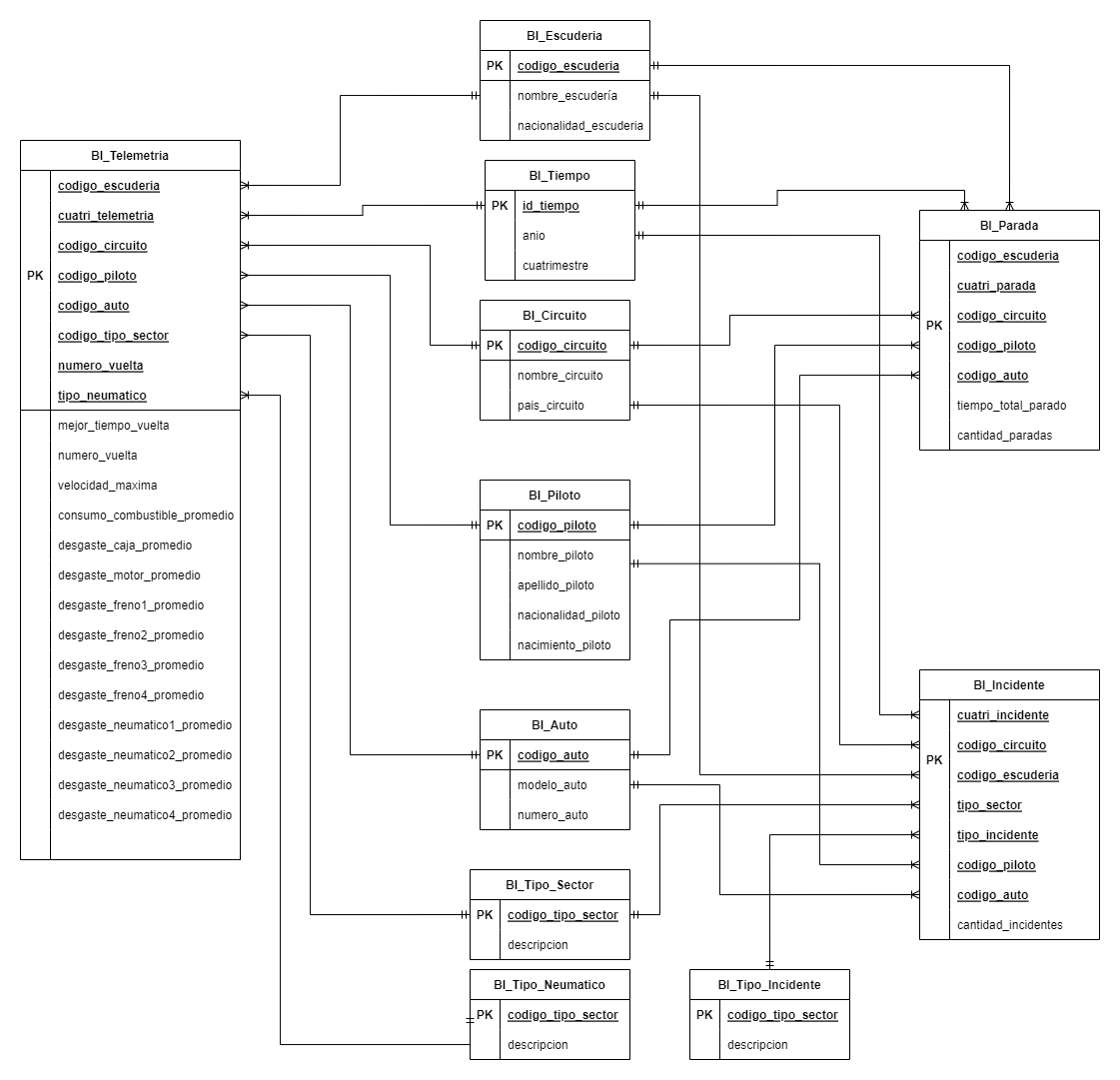
Optamos por cargar cada tabla por separado y de manera individual, respetando el orden que nos determinen las relaciones entre las distintas entidades de nuestro modelo. Primero cargamos las tablas que no referenciaban a ninguna otra (circuito, escudería, motor, neumático, freno, caja) y de ahí continuamos migrando el resto de las tablas, que dependían de las anteriores ya creadas.

A lo largo de este proceso nos encontramos con que había algunos atributos de nuestras tablas que no pertenecían a la tabla Maestra, por esta razón es que decidimos hacer un LEFT JOIN para asociar dichos atributos y completar la carga de la tabla.

En particular, para la carga de las tablas Freno, Telemetría Freno, Neumático, Telemetría Neumático y Cambio de Neumático decidimos realizar un insert distinto para cada freno y neumático ya que consideramos necesario cargarlos por separado.

Por ejemplo, para la carga de la tabla Telemetria\_Freno hicimos 4 inserts distintos para cada freno en particular (Tele\_Freno1, Tele\_Freno2, Tele\_Freno3, Tele\_Freno4).

# Modelo de datos BI:



# Estrategia de migración del modelo BI:

Comenzamos por cargar las tablas de dimensiones que en su mayoría resultaron ser inserciones planas desde las tablas del modelo transaccional. Para la dimensión del tiempo (anio, cuatrimestre) tuvimos que implementar y utilizar una función que determina el número de cuatrimestre concreto a partir de una fecha dada.

Determinamos que deben existir tres tablas de hechos: La de telemetrías, la de paradas en box y la de incidentes.

Cada una de las tablas de hecho está relacionada con todas las dimensiones posibles (siempre y cuando tengan sentido, por ejemplo la dimensión tipo de incidente solo aplica a la tabla de incidentes) para lograr la mayor granularidad alcanzable de la información de los hechos que luego serán consultados en las vistas.

Entendemos que si una vista requiere cierta información de hechos a partir de menos dimensiones de las que cuenta la tabla de hechos, esta información simplemente se obtiene haciendo agregaciones sobre las dimensiones “no necesitadas” (sea SUM, MAX, AVG, etc… dependiendo de la vista). De esta manera, sigue existiendo la posibilidad de agrupar los hechos de otras formas así lo necesitase algún requisito futuro a través de más vistas.

Por esto, las columnas de hechos fueron pensadas a partir de los requisitos de las vistas, pero no son directamente lo que piden las vistas, sino información más granular que luego es agregada según cada caso.

## Algunas Aclaraciones Puntuales:

Sobre la vista de menor tiempo de vuelta por escudería:

* Detectamos que al consultar la vista todos los mejores tiempos de vuelta eran 0.

Inicialmente creímos que fue un error en la lógica de la carga de los datos, pero luego haciendo revisiones sobre la tabla maestra, vimos que hay casos puntuales de autos (de hecho, todos aquellos con el numero 2) donde en la 3er vuelta de cada carrera solo contaban con una única telemetría en el instante cero.

Como el tiempo de vuelta de un auto se calcula a partir de tomar el último tiempo de telemetría dentro de una misma vuelta, todos aquellos autos en esta situación resultan en tiempos de vuelta calculados = 0, y se convierten en un elemento absorbente a la hora de buscar mínimos.

Podríamos haber realizado algún tipo de sanitización sobre estos datos, pero como no había indicaciones concretas sobre como abordarlo, decidimos no hacerlo, pero aclarando por las dudas de la existencia de los mismos.

Sobre la vista de desgastes promedio:

* Decidimos dejar por separados los componentes (neumáticos y frenos 1 a 4) para tener información más específica sobre el comportamiento de las componentes (ya que corresponden a diferentes secciones físicas del auto, y podrían no todos verse afectados de igual manera).