

11-1-2019

Cuadro de mando de sector
ferroviario
TFG Grado Ingeniería Informática

Memoria final



Rubén Franco Villa
Grado de Ingeniería Informática
Business Intelligence

Humberto Andrés Sanz
Atanasi Daradoumis Haralabus

Enero 2019



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

FICHA DEL TRABAJO FINAL

Título del trabajo:	Cuadro de mando integral de una empresa del sector ferroviario
Nombre del autor:	Rubén Franco Villa
Nombre del consultor:	Atanasi Daradoumis Haralabus
Fecha de entrega (mm/aaaa):	01/2019
Área del Trabajo Final:	Bussiness Intelligence
Titulación:	<i>Grado de Ingeniería Informática</i>

Resumen del Trabajo (máximo 250 palabras):

El primer ferrocarril fue construido en el siglo XIV. Desde entonces, el sector ha evolucionado mucho. Existe una gran complejidad tecnológica en un tren o un tranvía. Uno de los grandes retos del siglo XXI es mejorar la eficiencia de los transportes públicos y la movilidad de los ciudadanos. Además, ello permitirá reducir el tráfico de las grandes ciudades y la emisión de dióxido de carbono. El sector ferroviario genera mucha información, cuyo análisis nos permite conocer los hábitos de la población y mejorar el servicio.

El objetivo del presente trabajo es establecer una serie de indicadores que nos muestren el estado del servicio que presta una empresa del sector para que, posteriormente, podamos tomar decisiones a la vista de dichos indicadores.

Para ello, desarrollaremos un cuadro de mando integral que represente los indicadores y el resultado óptimo.

Abstract (in English, 250 words or less):

The first railway was built in the 14th century. Since then, the sector has evolved a lot. Nowadays there is a great technological complexity in a train or a tram. One of the great challenges of the 21st century is to improve the efficiency of public transport and the mobility of citizens. In addition, this will reduce traffic in large cities and carbon dioxide emissions. The railway sector generates a lot of information, whose analysis allows us to know the habits of the population and improve the service.

The objective of this work is to establish a series of indicators that show us the state of the service provided by a company in the sector and so we can make decisions in the light of these indicators.

To do this, we'll develop a comprehensive scorecard that represents the indicators and the optimal result.

Palabras clave (entre 4 y 8):

Eficiencia, transporte, Cuadro de mando integral, indicadores

Índice

Control de Cambios.....	1
1. Introducción	2
Contexto y justificación del Trabajo	2
1.1. Objetivos del Trabajo.....	2
1.2. Enfoque y método seguido	3
1.3. Planificación del Trabajo.....	3
1.4. Breve sumario de productos obtenidos	4
1.5. Breve descripción de los otros capítulos de la memoria	4
2. Implementación del proyecto	5
2.1. Fase análisis.....	5
2.1.1. Base de datos.....	5
2.1.2. Carga de datos	5
2.1.3. Generación de informes	7
2.1.4. Indicadores.....	7
2.1.5. Diagrama funcional.....	13
2.2. Fase de diseño	14
2.2.1. Arquitectura	14
2.2.2. Base de datos.....	15
2.2.3. Extracción transformación y carga de datos.....	18
2.2.4. Indicadores.....	19
2.2.5. Cuadro de mando integrado.....	20
2.3. Fase de desarrollo	23
2.3.1. Preparación entorno de trabajo	23
2.3.2. Base de datos.....	27
2.3.3. Creación ETLs.....	30
2.3.4. Programación ETLs.....	37
2.3.5. Creación CMI.....	40
3. Conclusiones	44
4. Líneas de futuro	44
5. Glosario	45
6. Bibliografía.....	45

Índice Tablas

Tabla 1 - Cronograma	3
Tabla 2 - Cuadro comparativo	7
Tabla 3 - Indicador D1	8
Tabla 4 - Indicador D2.....	8
Tabla 5 - Indicador D3.....	9
Tabla 6 - Indicador D4.....	10
Tabla 7 - Indicador D5.....	10
Tabla 8 - Indicador D6.....	11
Tabla 9 - Indicador D7	11
Tabla 10 - Indicador D8.....	12
Tabla 11 - Indicador D9.....	12
Tabla 12 - Indicador D10.....	13
Tabla 13 - SIMOVE	16
Tabla 14 - Cancelaciones.....	17
Tabla 15 - Ventas	17
Tabla 16 - Indicadores.....	17
Tabla 17 - Todos los indicadores	20
Tabla 18 – Glosario	45

Índice Ilustraciones

Ilustración 1 - Diagrama de flujo.....	14
Ilustración 2 - Arquitectura solución	15
Ilustración 3 – Tablas CMI.....	16
Ilustración 4 – CMI Indicadores	20
Ilustración 5 – CMI Ventas	21
Ilustración 6 – CMI Validaciones	22
Ilustración 7 – CMI Viajes.....	22
Ilustración 8 – CMI Mantenimiento	23
Ilustración 9 – Configuración VirtualBox.....	24
Ilustración 10 – Configuración Red VirtualBox	25
Ilustración 11 – Servicios SQL	25
Ilustración 12 – Conexión SQL Server	26
Ilustración 13 – Talend	27
Ilustración 14 – Microstrategy Desktop	27
Ilustración 15 – Base de Datos CMI	29
Ilustración 16 – Proyecto Talend TFG.....	30
Ilustración 17 – Trabajo Main	30
Ilustración 18 – Trabajo SIMOVE	30
Ilustración 19 – Trabajo Validaciones.....	31
Ilustración 20 – Trabajo Ventas.....	31
Ilustración 21 – Trabajo Indicadores	32
Ilustración 22 – Conexión SQL server CMI	32
Ilustración 23 – Conexión Oracle VIAMOVIL.....	32
Ilustración 24 – Input Oracle SIMOVE.....	33
Ilustración 25 – Output SQL Server CMI	33

Ilustración 26 – tMap	34
Ilustración 27– Ejecución Job.....	34
Ilustración 28– Datos tabla Indicadores	35
Ilustración 29 – Datos tabla Simove	35
Ilustración 30 – Datos tabla Validaciones.....	36
Ilustración 31 – Datos tabla Ventas	37
Ilustración 32 – Exportar trabajo Talend.....	38
Ilustración 33 – Script trabajo Talend	38
Ilustración 34 – Creación de tarea parte 1	39
Ilustración 35 – Creación de tarea parte 2	39
Ilustración 36 – Creación de tarea parte 3	40
Ilustración 37 – Dossier CMI	40
Ilustración 38 – Conjuntos de datos	41
Ilustración 39 – Ejemplo de conjunto de datos	41
Ilustración 40 – Informe de indicadores.....	41
Ilustración 41 – Informe de ventas Diarias	42
Ilustración 42 – Informe de viajes diarios	42
Ilustración 43 – Informe de velocidad por conductor	43
Ilustración 44 – Informe de desgaste de las ruedas de los vehículos	43

Control de Cambios

Versión	Fecha	Autor	Revisor	Descripción
00.01	18/11/2018	Rubén Franco	Humberto Andrés	
00.02	11/12/2018	Rubén Franco	Humberto Andrés	
00.03	11/01/2019	Rubén Franco		

1. Introducción

Contexto y justificación del Trabajo

Constituida el 22 de enero de 2001, la empresa Metropolitano de Tenerife (a partir de ahora MTSA) es una Sociedad Anónima. De capital social 100% propiedad del Cabildo de Tenerife desde el 4 de abril de 2017, día en el que la Corporación insular formalizó la compra del 14% de Tenemetro, S.L. (Transdev/Sacyr/Ineco). Originariamente, MTSA nace para poner en marcha un transporte alternativo en forma de líneas ferroviarias en la isla de Tenerife. Para ello: el primer objetivo de MTSA es realizar todos los estudios necesarios para el diseño de una Red Ferroviaria eficaz y eficiente, dentro del ámbito de actuación legislativo y normativo de la Comunidad Autónoma de Canarias, España y la Unión Europea. El segundo objetivo, es la construcción y puesta en marcha de las primeras líneas ferroviarias que contribuyan a aliviar la congestión circulatoria, facilitar y satisfacer las demandas de movilidad de la población de Tenerife y sus visitantes. MTSA es miembro de la UITP, “Union Internationale des Transports Publics” y es uno de los dos representantes españoles que participan como miembros en su Comité de Metros Ligeros, y pertenece a la Asociación Latinoamericana de Metros y Subterráneos, Alamys, y a la asociación de empresas gestoras de los transportes urbanos colectivos, ATUC. La organización de la empresa, está constituida de varios departamentos, entre los que se encuentra el de Informática e Ingeniería de Sistemas (a partir de ahora IeIS). Dicho departamento no dispone de un CIO, ni de un área de Business Intelligence (a partir de ahora BI) para explotar la información que genera la empresa. Los distintos departamentos de la empresa generan gran cantidad de información, que únicamente explotan a través de informes propios de la aplicación. Lo que dificulta la sinergia entre los distintos sistemas y la creación de indicadores.

1.1. Objetivos del Trabajo

Demostrar como la aplicación de una herramienta de BI permite gestionar la información para conseguir mejoras en la toma de decisiones de MTSA.

El principal objetivo del trabajo fin de grado (a partir de ahora TFG), es implementar un cuadro de mando integral (a partir de ahora CMI), en el cual se visualicen una serie de indicadores ya definidos previamente, que muestren el estado de los sistemas más críticos de la empresa. Para ello, se deberá diseñar una base de datos nueva, donde almacenar la información de los distintos sistemas, mediante un proceso de extracción, transformación y carga (a partir de ahora ETL) de los datos. También, se deberán definir indicadores y los distintos umbrales de los mismos. Finalmente, se creará un CMI que represente el estado de los procesos críticos de la empresa de una forma clara y amigable para poder establecer conclusiones.

1.2. Enfoque y método seguido

He decidido utilizar una metodología ágil (Scrum), debido a que el proyecto tiene un alcance y tiempo limitado, lo que unido a los posibles cambios que proponga el director del proyecto, hace necesario una metodología flexible y efectiva. Consiguiendo así que el impacto de los cambios sea el menor posible y no afecte a la calidad del producto final. Por lo tanto, será clave la comunicación con el director del TFG.

1.3. Planificación del Trabajo

Tarea	Duración	Inicio	Fin
1. Trabajo Fin de Grado	85 días	19/09/2018	13/01/2019
1.1. Definición del Proyecto	4 días	19/09/2018	24/09/2018
1.2. Plan de Trabajo	15 días	25/09/2018	13/10/2018
Hito PEC1	0 días	15/10/2018	15/10/2018
1.3. Requisitos	4 días	16/10/2018	19/10/2018
1.3.1. Definir indicadores	4 días	16/10/2018	19/10/2018
1.4. Análisis	15 días	22/10/2018	09/11/2018
1.4.1. Análisis de los requerimientos	2 días	22/10/2018	23/10/2018
1.4.2. Análisis de los orígenes de datos	6 días	24/10/2018	31/10/2018
1.4.3. Análisis del IHM	3 días	01/11/2018	05/11/2018
1.4.4. Análisis herramientas a usar	4 días	06/11/2018 9:00	09/11/2018
1.5. Diseño	12 días	12/11/2018 9:00	27/11/2018
1.5.1. Diseño de la Base de Datos	3 días	12/11/2018 9:00	14/11/2018
1.5.2. Diseño de los ETLs	2 días	15/11/2018 9:00	16/11/2018
1.5.3. Diseño de los indicadores	2 días	19/11/2018 9:00	20/11/2018
Hito PEC2	0 días	19/11/2018 9:00	19/11/2018
1.5.4. Diseño del CMI	5 días	21/11/2018	27/11/2018
1.6. Implementación	27 días	28/11/2018	03/01/2019
1.6.1 Preparación entorno de trabajo	3 días	28/11/2018	30/11/2018
1.6.2. Creación de base de datos	2 días	03/12/2018	04/12/2018
1.6.3. Creación ETLs	7 días	05/12/2018	13/12/2018
1.6.4. Programación ETLs	1 día	14/12/2018	14/12/2018
1.6.5. Creación CMI	14 días	17/12/2018	03/01/2019
Hito PEC3	0 días	12/12/2018	12/12/2018
1.7. Pruebas	5 días	04/01/2019	10/01/2019
1.8. Cierre Proyecto	2 días	11/01/2019	13/01/2019
Entrega Final	0 días	13/01/2019	13/01/2019

Tabla 1 - Cronograma

1.4. Breve resumen de productos obtenidos

Los distintos entregables que se darán al finalizar el proyecto son:

- Memoria del Proyecto. Este es el entregable principal, donde se define todo el trabajo realizado durante la TFG.
- Presentación en formato video que explicará la TFG.
- Presentación en Powerpoint (o producto similar) que usaré en el video.

1.5. Breve descripción de los otros capítulos de la memoria

Lo que describe a continuación el presente documento, es como se ha elaborado las principales fases del TFG.

- Fase de análisis
- Fase de diseño
- Fase de implementación

Estas fases se han dividido en subapartados para una mayor comprensión.

2. Implementación del proyecto

2.1. Fase análisis

2.1.1. Base de datos

Para este trabajo, será necesario crear una base de datos en local donde almacenar la información que quiero explotar de los distintos orígenes de datos. Para ello, he analizado distintas alternativas como son:

- Postgresql
- SQL Server Express
- MariaDB
- MySQL

Al final, me he decantado por usar un SQL Server 2008 Express, el motivo es principalmente la curva de aprendizaje, al ser un tiempo muy limitado el plazo de entrega trabajo, he decidido ir a la base de datos que más conozco. Para administrarla usaré el SQL Server Studio Management.

2.1.2. Carga de datos

El sistema de Explotación de Información mostrará información proveniente de los diferentes sistemas de origen.

- SIMOVE → Oracle 11 6 (sistema de operaciones)
- Ventas → Oracle 11 6 (sistema de billeteaje)
- Cancelaciones → Oracle 11 6 (sistema de billeteaje)
- Indicadores → SQL Server 2005 (sistema Intranet)

Para ello, se realizará desde los mismos una carga de datos hacia el modelo de datos de la plataforma informacional, consistente en:

- Capturar los datos de los distintos entornos e incorporarlos a unos soportes “estándar” de intercambio.
- Homogeneizar los datos capturados: reestructurar campos, quitar ciertos datos no significativos, traducir formatos de campos, comprobar consistencia e integridad de contenidos y detectar errores.
- Enriquecer el contenido analítico de la información: calcular campos derivados, mezclar y operar sobre datos de distintas fuentes, etc.

Asegurar la calidad de los datos: es absolutamente fundamental que la información que se carga en el Sistema esté suficientemente validada y no ofrezca ninguna duda sobre su fiabilidad.

La información se cargará en el modelo de datos del Cuadro de Mando Integral de forma periódica. Este periodo se determina con el objeto de mantener

actualizados los datos respecto a los existentes en los sistemas origen, y teniendo en cuenta:

- La periodicidad de cálculo de los indicadores afectados por la carga
- El tipo de información.
- El volumen de información.

En definitiva, se ajusta la ventana temporal de extracción y carga de información a los requerimientos funcionales. El objetivo es armonizar los tiempos de extracción y carga de datos de acuerdo con las restricciones de disponibilidad de los entornos origen y destino y las propias necesidades del usuario final.

El proceso de carga de información será diario e incremental, añadiéndose desde el origen la información nueva y/o modificada. Por lo tanto, cada día se cargará la información nueva y/o modificada de las tablas que tienen datos que afecten a indicadores con periodicidad de cálculo diario. Lo mismo sucederá para los indicadores con periodicidad de cálculo semanal y mensual.

En el caso de los registros de orígenes de datos de los que no se tenga la capacidad de comprobar su modificación o creación, serán cargados en su totalidad en cada proceso de carga.

Para ello, he analizado 2 herramientas:

- Talend
- Integration Services Microsoft

Cuadro comparativo:

Software	SSIS	Talend
Desarrollador	Microsoft	Talend
Objetivo	ETL	ETL
Pros	<ul style="list-style-type: none"> • Permite ejecutar muchos procesos en paralelo. • Muchas herramientas para transformar datos durante el proceso de migración. 	<ul style="list-style-type: none"> • Interfaz fácil de usar • Diseñar trabajos es muy sencillo
Contras	<ul style="list-style-type: none"> • No puede copiar y pegar anotaciones, por lo que, si tiene varias similares, debe escribir cada una de ellas desde cero. • Problemas en conversión de datos 	<ul style="list-style-type: none"> • Sincronizar con Git es un poco difícil. • Puede enfrentar problemas con ciertos problemas de big data.

Coste	• Elevado	• Gratuito
Soporte	• Excelente	• Excelente

Tabla 2 - Cuadro comparativo

Me he decantado por el Talend versión 6.0.1, por sus buenas prestaciones, soporte, usabilidad y que al tratarse de un trabajo académico valoro positivamente que sea gratuito.

2.1.3. Generación de informes

Para los informes incluidos en el CMI he analizado distintas alternativas:

- QlikView
- Microstrategy
- Power BI
- SAP BI

Al final, he decidido usar Microstrategy debido a que existe una versión denominada Microstrategy Desktop que es gratuita, la cual cumple con creces las necesidades del proyecto.

2.1.4. Indicadores

Para determinar lo que queremos analizar de la empresa, se van a crear una serie de indicadores/KPIs, que nos indicaran el grado de cumplimiento de los distintos aspectos analizados respecto al umbral establecido.

Para ello usaremos unos indicadores ya previamente definidos por el ECIT y suministrada la información por los diferentes sistemas, que abordan las principales áreas de la empresa.

Los indicadores se calculan por tanto de forma automática con los datos que desprende los orígenes de datos.

Los indicadores tienen un umbral mínimo a alcanzar, ya que el objetivo de los mismos es que la calidad del servicio no decaiga por debajo de esos umbrales. En el supuesto que en un intervalo dado no alcance el umbral mínimo, dicho umbral penalizará en las políticas que aplica el Excelentísimo Cabildo Insular de Tenerife (a partir de ahora ECIT) en forma de subvenciones al transporte.

A continuación, analizaré como se han calculado los indicadores D1 - D12, ya que son indicadores que nos vienen precalculados.

2.1.4.1. Indicador D1

La fiabilidad evalúa el cumplimiento de los servicios programados. Es decir, analiza la diferencia entre los viajes programados efectivamente realizados y el total de los programados.

D1 – VIAJES PERDIDOS	
Objetivo	Medir el grado de cumplimiento de los servicios programados
Forma de cálculo	$FHR = \frac{NVc}{NVp} * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NVc: Número de viajes comerciales completos efectuados. • NVp: Número de viajes programados para la semana.
Origen de Datos	SAE
Rango	< 96 → Rojo, >= 96 y < 98 Ámbar, >= 98 Verde.
Tipo de indicador	Velocímetro
Filtros	Viajes comerciales que afectan al ECIT

Tabla 3 - Indicador D1

2.1.4.2. Indicador D2

Evalúa el grado de cumplimiento de los servicios realizados frente a los programados, considerando no realizados los que han llegado a su destino con un retraso superior a 5 minutos.

D2 – VIAJES RETRASADOS	
Objetivo	Medir el grado de cumplimiento de los viajes comerciales completos
Forma de cálculo	$FSr = \frac{NVb}{NVs} * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NVb: Número de viajes comerciales completos efectuados o el periodo de evaluación para los cuales el retraso en la llegada a la parada final no supere los 5 minutos frente al horario programado. • NVs: Número de viajes programados para el semestre o periodo de evaluación, es decir, todos aquellos que tienen un tiempo teórico.
Origen de Datos	SAE
Rango	< 96 → Rojo, >= 96 y < 98 Ámbar, >= 98 Verde.
Tipo de indicador	Velocímetro
Filtros	Viajes comerciales que afectan al ECIT

Tabla 4 - Indicador D2

2.1.4.3. Indicador D3

Evalúa el grado de puntualidad de las salidas de los servicios realizados.

D3 – VIAJES ADELANTADOS	
Objetivo	Medir el grado de cumplimiento de las salidas de los servicios
Forma de cálculo	$PSa(\text{diario}) = (NVe - NVa) / NVe * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NVe: Número de viajes comerciales completos efectuados durante el día. • NVa: Número de viajes comerciales tempranos efectuados durante el día.
Origen de Datos	SAE
Rango	< 96 → Rojo, >= 96 y < 98 Ámbar, >= 98 Verde.
Tipo de indicador	Velocímetro
Filtros	Viajes comerciales que afectan al ECIT

Tabla 5 - Indicador D3

2.1.4.4. Indicador D4

Evalúa el grado de puntualidad de las salidas tardías.

D4 – VIAJES SALIDA TEMPRANA	
Objetivo	Medir el grado de cumplimiento de las salidas de los servicios
Forma de cálculo	$PSt(\text{diario}) = (NVe - NVt) / NVe * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NVe: Número de viajes comerciales completos durante el día. • NVt: Número de viajes comerciales tardíos efectuados durante el día, no se considera que sale tardío si hay menos de 5 minutos desde la llegada del viaje anterior hasta la salida de este, aunque salga con 5 minutos de retraso respecto al planificado.
Origen de Datos	SAE

Rango	< 96 → Rojo, >= 96 y < 98 Ámbar, >= 98 Verde.
Tipo de indicador	Velocímetro
Filtros	Viajes comerciales que afectan al ECIT

Tabla 6 - Indicador D4

2.1.4.5. Indicador D5

Se define como el cálculo de la puntualidad en los primeros y últimos viajes comerciales.

D5 – VIAJES SALIDA TARDÍA	
Objetivo	Medir el grado de puntualidad de los primeros y últimos viajes
Forma de cálculo	$PS_x = \frac{(NPP_p + NUP_p) - (NPP_e + NUP_e)}{(NPP_p + NUP_p)} * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NPP_p: Número de primeros viajes comerciales programados durante el periodo. • NPP_e: Número de primeros viajes comerciales efectuados en las que el tiempo de salida de la primera parada es inferior a 15 segundos respecto al tiempo teórico. • NUP_p: Número de últimos viajes comerciales programados durante el periodo. • NUP_e: Número de últimos viajes comerciales efectuados en las que el tiempo de salida de la primera parada es inferior a 15 segundos respecto al tiempo teórico.
Origen de Datos	SAE
Rango	< 96 → Rojo, >= 96 y < 98 Ámbar, >= 98 Verde.
Tipo de indicador	Velocímetro
Filtros	Viajes comerciales que afectan al ECIT

Tabla 7 - Indicador D5

2.1.4.6. Indicador D6

Porcentaje de las limpiezas que se han realizado correctamente frente al total de las realizadas, se consideran no realizadas correctamente las que tienen como puntuación 9 o más puntos.

D6 - LIMPIEZA DE LOS TRANVÍAS	
Objetivo	Control de la limpieza de los tranvías
Forma de cálculo	$LT = \frac{NLTe}{NLTp} * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NLTe: Número de limpiezas realizadas correctamente • NLTp: Número de limpiezas realizadas
Origen de Datos	GMAO
Rango	< 98 → Rojo, >= 98 y < 99 Ámbar, >= 99 Verde.
Tipo de indicador	Velocímetro

Tabla 8 - Indicador D6

2.1.4.7. Indicador D7

Porcentaje de las reparaciones resueltas en 48 horas frente al total. Este periodo de 48 horas contempla desde la detección del problema hasta la terminación del último trabajo realizado en la orden de trabajo.

D7 – REPARACIÓN DE DAÑOS EN LOS TRANVÍAS	
Objetivo	Control del tiempo de respuesta en averías.
Forma de cálculo	$IT = \frac{NITc - NITn}{NITc} * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NITc: Número de OTs relacionadas con daños en los vehículos creadas en el periodo de evaluación. • NITn: Número de OTs relacionadas con daños en los vehículos creadas en el periodo de evaluación y no solucionadas en 48 horas naturales.
Origen de Datos	GMAO
Rango	< 95 → Rojo, >= 95 y < 98 Ámbar, >= 98 Verde.
Tipo de indicador	Velocímetro

Tabla 9 - Indicador D7

2.1.4.8. Indicador D8

Porcentaje de las limpiezas que se han realizado correctamente frente al total de las realizadas, se consideran no realizadas correctamente las que tienen como puntuación 9 o más

D8 – LIMPIEZA DE LAS PARADAS	
Objetivo	Control de la limpieza de las paradas
Forma de cálculo	$LT = \frac{NLTe}{NLTp} * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NLTe: Número de limpiezas realizadas correctamente • NLTp: Número de limpiezas realizadas
Rango	< 98 → Rojo, >= 98 y < 99 Ámbar, >= 99 Verde.
Tipo de indicador	Velocímetro

Tabla 10 - Indicador D8

2.1.4.9. Indicador D9

Porcentaje de las reparaciones resueltas en 48 horas frente al total. Este periodo de 48 horas contempla desde la detección del problema hasta la terminación del último trabajo realizado en la OT.

D9 – REPARACIÓN DE DAÑOS EN LAS PARADAS	
Objetivo	Control del tiempo de respuesta en averías.
Forma de cálculo	$IT = \frac{NITc - NITn}{NITc} * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • NITc: Número de OTs relacionadas con daños en las paradas creadas en el periodo de evaluación. • NITn: Número de OTs relacionadas con daños en las paradas creadas en el periodo de evaluación y no solucionadas en 48 horas naturales.
Origen de Datos	GMAO
Rango	< 95 → Rojo, >= 95 y < 98 Ámbar, >= 98 Verde.
Tipo de indicador	Velocímetro

Tabla 11 - Indicador D9

2.1.4.10. Indicador D10

Porcentaje de disponibilidad de las expendedoras. Las expendedoras generan diferentes avisos y en función de estos puede considerarse que están fuera de servicio (cofre de monedas lleno, cofre de billetes bancarios llenos, puerta

abierta, etc....). Inicialmente se considerará una expendedora fuera de servicio cuando emite el aviso de fuera de servicio.

D10 – DISPONIBILIDAD DE LAS MAQUINAS EXPENDEDORAS	
Objetivo	Medir la disponibilidad de las Expendedoras
Forma de cálculo	$DB = \frac{Tsc - Tindis}{Tsc} * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • Tsc: Tiempo de Servicio Comercial: Tiempo de servicio con viajes comerciales (en minutos) x Número de Expendedoras. • Tindis: Tiempo de Indisponibilidad de las Expendedoras observado en el periodo en minutos.
Rango	< 98 → Rojo, >= 98 y < 98 Ámbar, >= 98 Verde.
Tipo de indicador	Velocímetro

Tabla 12 - Indicador D10

2.1.5. Diagrama funcional

A través del siguiente esquema se presenta la funcionalidad genérica que el sistema aportará al usuario, y que será detallada en los sucesivos apartados. Básicamente, la aplicación, tras una correcta extracción y carga de datos desde los sistemas origen, ofrecerá al usuario la posibilidad de visualizar los datos a distintos niveles:

- Nivel elemental (Ej.: Viaje)
- Nivel Global (Indicadores ECIT)

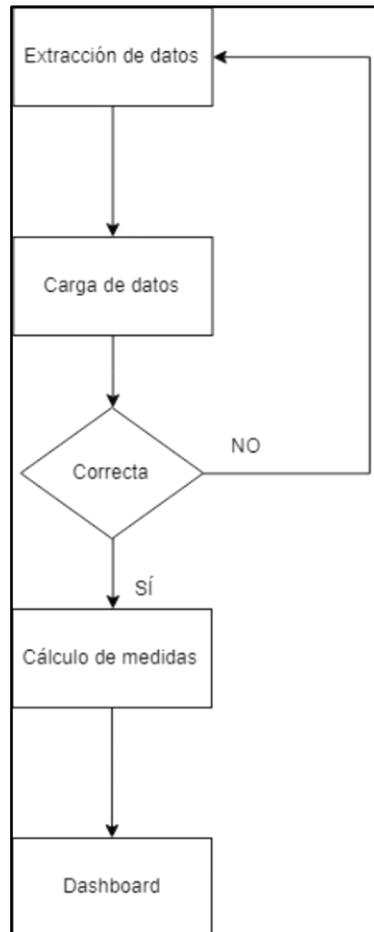


Ilustración 1 - Diagrama de flujo

2.2. Fase de diseño

En este apartado vamos a definir en mayor detalle los requisitos, basándonos en el análisis previo que se ha hecho.

2.2.1. Arquitectura

En el presente apartado se muestra en forma de diagrama los distintos sistemas involucrados.

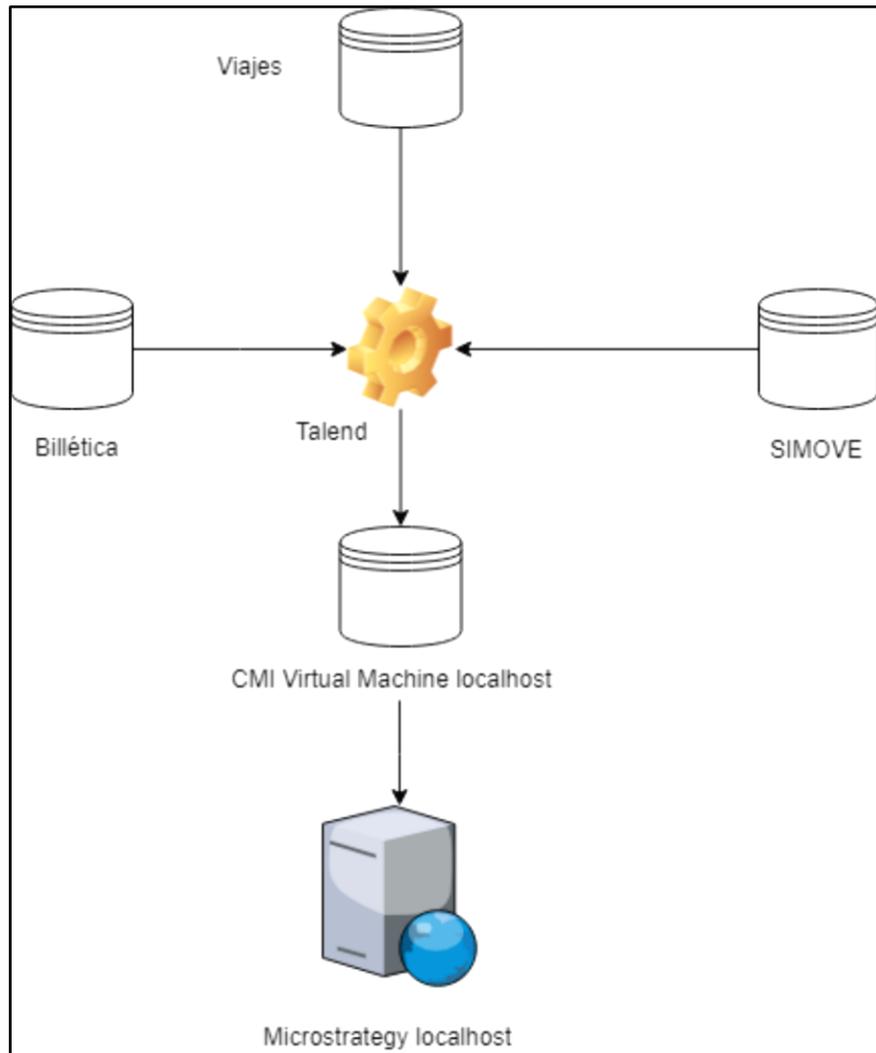


Ilustración 2 - Arquitectura solución

2.2.2. Base de datos

Para ello, usaré un modelo entidad-relación que se implementará en SQL Server 2008 Express. A continuación, definiremos las distintas entidades que formarán parte de la base de datos integrada CMI.

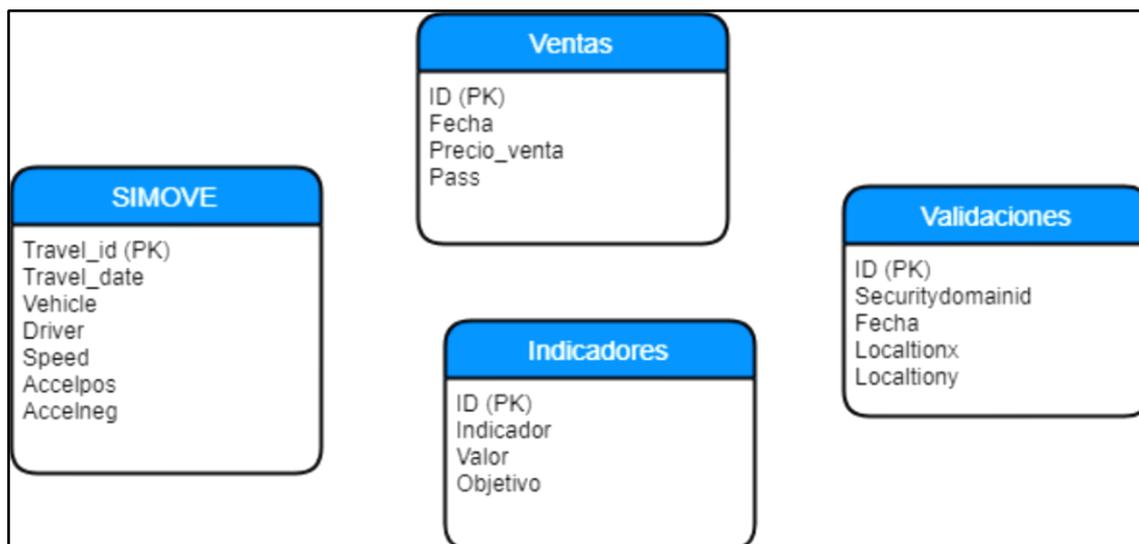


Ilustración 3 – Tablas CMI

Tabla SIMOVE

En esta tabla se almacenarán los viajes realizados por cada conductor, cada día, y sus estadísticas de velocidad y aceleración.

Campo	Tipo de dato	Definición
Travel_id (PK)	Nvarchar(20)	Identificador de cada viaje que realiza un tranvía
Travel_date	Date	Fecha del viaje en formato dd/mm/yyyy
Vehicle	Nvarchar(10)	Identificador del vehículo
Driver	Nvarchar(20)	Identificador del conductor
Speed	Float	Velocidad media
Accelpos	Float	Aceleración media positiva
Accelneg	Float	Aceleración media negativa

Tabla 13 - SIMOVE

Tabla Validaciones

En esta tabla almacenaremos las validaciones/cancelaciones de los usuarios. Es decir, los viajes que realizan los usuarios por fecha y geolocalización en caso de que está disponible.

Campo	Tipo de dato	Definición
ID (PK)	Integer	Identificador de cada viaje de los usuarios.

Securitydomainid	Bigint	Identificador del usuario
Fecha	Date	Fecha en la que se produce el viaje
Localtionx	Float	Coordenada x de donde se produjo el viaje
Localtiony	Float	Coordenada y de donde se produjo el viaje

Tabla 14 - Cancelaciones

Tabla de Ventas

En esta tabla se almacenarán la recaudación que se ha producido en el sistema agrupada por día y tipo de título (abono mensual, bono de 15€, etc...)

Campo	Tipo de dato	Definición
ID (PK)	Integer	Identificador auto-numérico de la tabla.
Fecha	Date	Fecha en la que se produce las ventas
Precio_venta	Float	Valor de la recaudación en Euros.
Pass	Nvarchar(500)	Nombre del título que el cliente ha comprado

Tabla 15 - Ventas

Tabla Indicadores

En esta tabla se almacenarán los valores obtenidos del cálculo de los principales indicadores (D1 – D10) y el umbral mínimo deseado.

Campo	Tipo de dato	Definición
ID (PK)	Integer	Identificador auto-numérico de la tabla.
Indicador	Nvarchar(500)	Descripción del indicador calculado
Valor	Float	Porcentaje obtenido del sistema.
Objetivo	Float	Umbral mínimo para cobrar las políticas del ECIT

Tabla 16 - Indicadores

2.2.3. Extracción transformación y carga de datos

Usaré la herramienta Talend, en la cual se realizarán los siguientes trabajos:

Trabajo Main



Trabajo SIMOVE



Trabajo Ventas



Trabajo Validaciones



Trabajo Indicadores



A continuación, se deberán crear las planificaciones necesarias para la ejecución automática de la carga de datos. Para ello, nos decantaremos por una tarea de Windows, ya que el Talend genera ficheros ejecutables que pueden ser fácilmente programables. Las ejecuciones serán todos los días a las 3 a.m. para así asegurarnos de no afectar el rendimiento del sistema operacional.

2.2.4. Indicadores

Los indicadores que se utilizarán para el cuadro de mando son los siguientes:

Indicador	Umbral	Origen	Descripción
D1 - Viajes Perdido	98%	Intranet	Ver apartado de análisis
D2 - Viajes Retrasados	98%	Intranet	Ver apartado de análisis
D3 - Viajes Adelantados	98%	Intranet	Ver apartado de análisis
D4 - Viajes Salida Temprana	98%	Intranet	Ver apartado de análisis
D5 - Viajes Salida Tardía	98%	Intranet	Ver apartado de análisis
D6 - Limpieza Vehículos	98%	Intranet	Ver apartado de análisis
D7 - Reparación Vehículos	98%	Intranet	Ver apartado de análisis
D8 - Limpieza Paradas	98%	Intranet	Ver apartado de análisis
D9 - Reparación Paradas	98%	Intranet	Ver apartado de análisis
D10 - Disponibilidad Expendedoras	98%	Intranet	Ver apartado de análisis
Velocidad Media	No aplica	SIMOVE	Media de las velocidades por conductor
Aceleración positiva media	No aplica	SIMOVE	Indicador de confort en las salidas de las paradas
Aceleración negativa media	No aplica	SIMOVE	Indicador de confort en las frenadas en las paradas
Diámetro rueda	No aplica	SIMOVE	Análisis del desgaste

			de la rueda
Total de viajes	No aplica	SIMOVE	Volumen de viajes realizados por los vehículos y conductores
Total de ventas	No aplica	Billética	Volumen de títulos vendidos
Recaudación	No aplica	Billética	Recaudación de la empresa
Total usuarios	No aplica	Billética	Usuarios que usan el transporte público
Total de validaciones	No aplica	Billética	Cantidad de viajes que realizan los usuarios

Tabla 17 - Todos los indicadores

2.2.5. Cuadro de mando integrado

Para el diseño de este apartado usaré una herramienta de diseño llamada Balsacmic Mockup, la cual me permitirá mostrar de una forma fácil el IHM. Todo lo aquí mostrado será desarrollado en la herramienta Microstrategy Desktop.

2.2.5.1. Indicadores

En el presente informe, podremos ver cuál es el grado de cumplimiento de los distintos indicadores del ECIT, así como un resumen de los viajes, usuarios y recaudación.

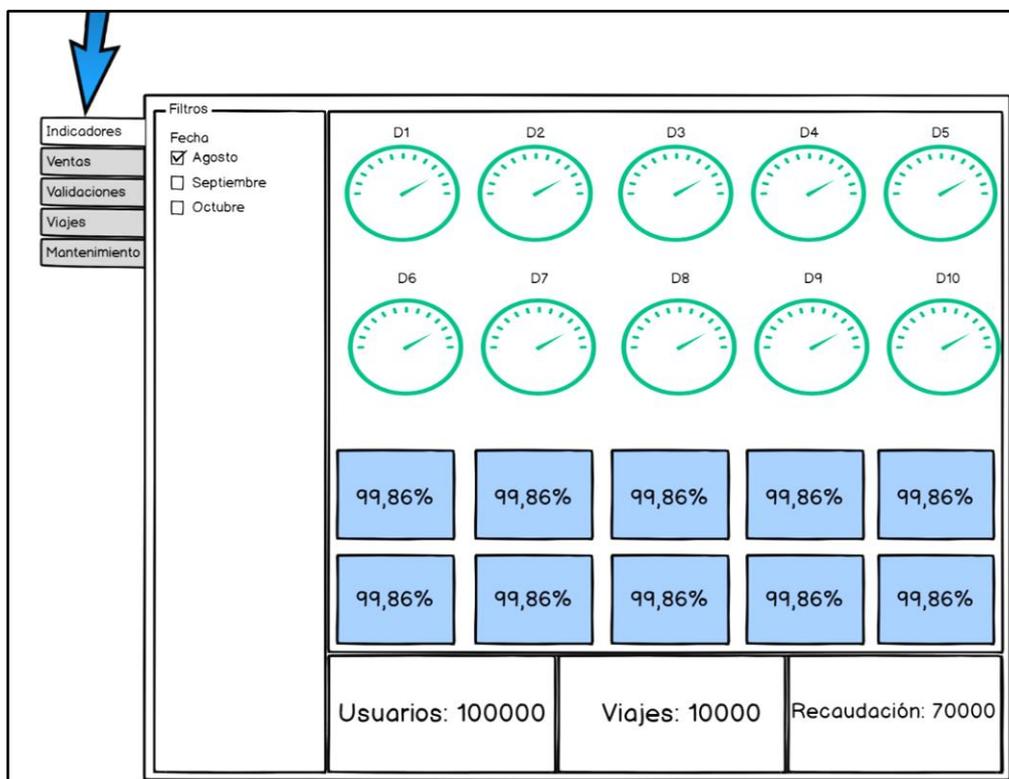


Ilustración 4 – CMI Indicadores

2.2.5.2. Ventas

En este informe podremos ver la evolución a lo largo del tiempo, de los distintos títulos, tanto en la recaudación como en el volumen de ventas.

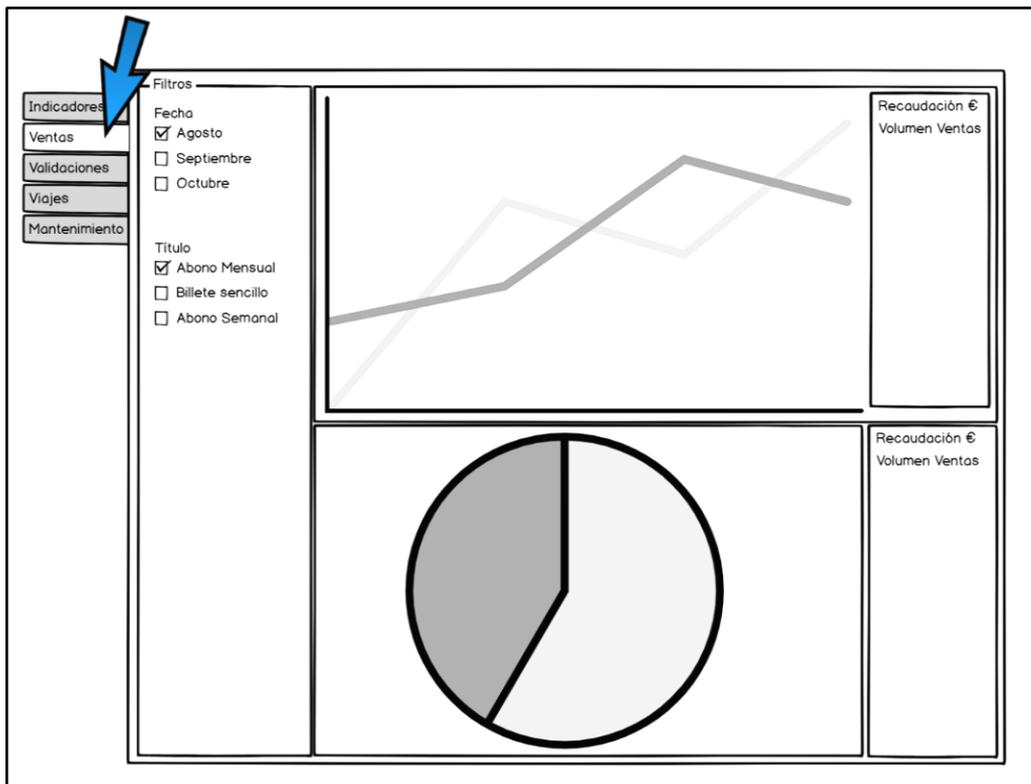


Ilustración 5 – CMI Ventas

2.2.5.3. Validaciones

En el capítulo de validaciones se podremos observar la evolución en el tiempo del volumen de validaciones, así como la localización de las mismas en un mapa.

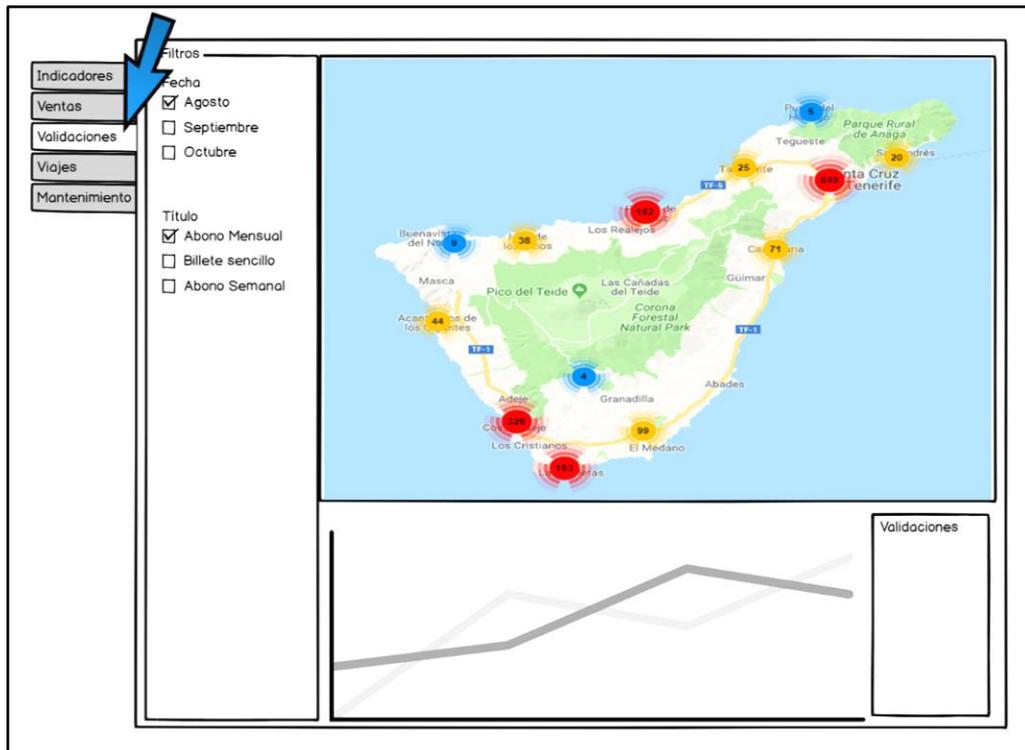


Ilustración 6 – CMI Validaciones

2.2.5.4. Viajes

El capítulo de viajes, debe mostrar los viajes comerciales que se han efectuado por los distintos vehículos y conductores a lo largo de los meses. También será necesario poder ver un ranking de velocidad comercial de conductores y el confort para el pasaje en estos viajes.

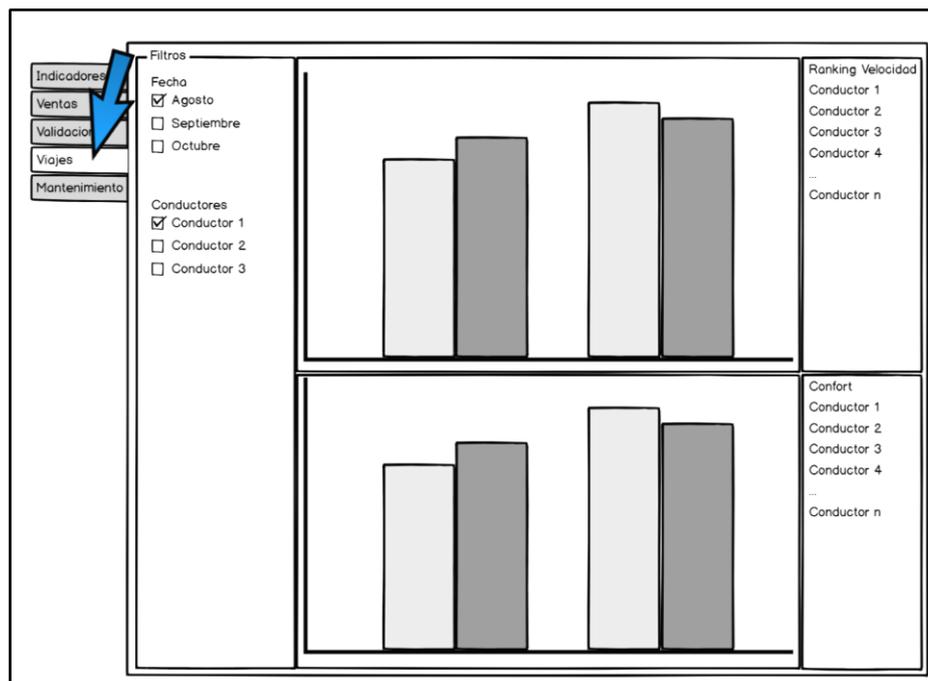


Ilustración 7 – CMI Viajes

2.2.5.5. Mantenimiento

En este apartado, se debe visualizar la evolución en forma de desgaste que sufre la rueda viaje a viaje, para poder hacer reperfilados o cambio de rueda de una forma más óptima.

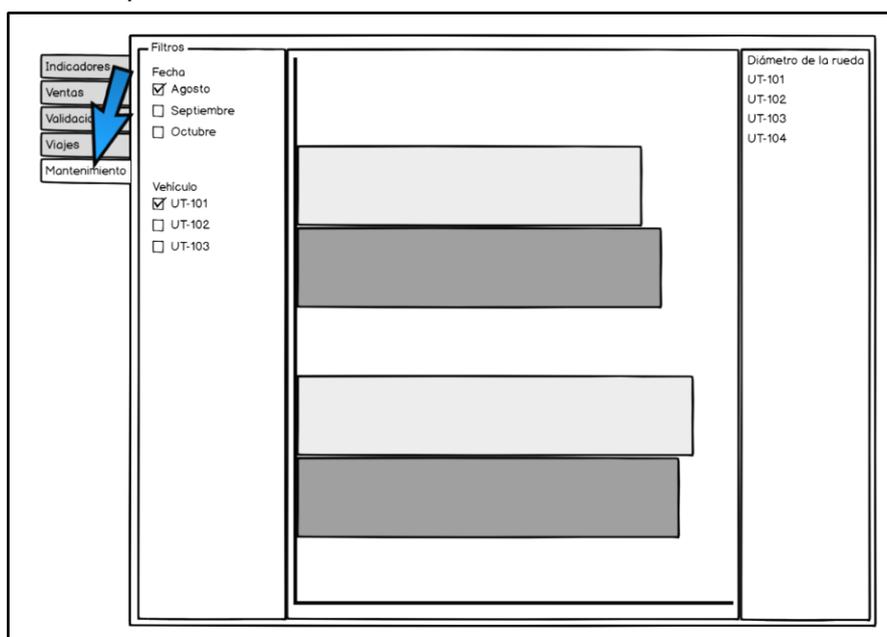


Ilustración 8 – CMI Mantenimiento

2.3. Fase de desarrollo

En el presente apartado, se detallará toda la fase de implementación del proyecto, desde la instalación de herramientas necesarias, hasta el desarrollo del cuadro de mando.

2.3.1. Preparación entorno de trabajo

En este apartado detallaré los pasos que he realizado para preparar el entorno de desarrollo.

2.3.1.1. Instalación de un servidor Virtual

Al no disponer de un servidor para desarrollar el TFG, he optado por mi ordenador personal. Para una mayor limpieza y realismo, he decidido instalar un servidor virtual con la herramienta gratuita VirtualBox.

El VirtualBox se puede descargar de <https://www.virtualbox.org/wiki/Downloads>.

Es necesaria una imagen ISO del sistema operativo que se desea utilizar. En mi caso, tengo varias que me facilita mi entorno de trabajo.

VirtualBox es una herramienta de la empresa Oracle que permite la virtualización de sistemas operativos “invitados”, dentro de un sistema operativo “anfitrión”. En mi caso el sistema operativo anfitrión es un Windows 10 Pro 64bit, y el invitado es un Windows Server 2008 R2 32bit.

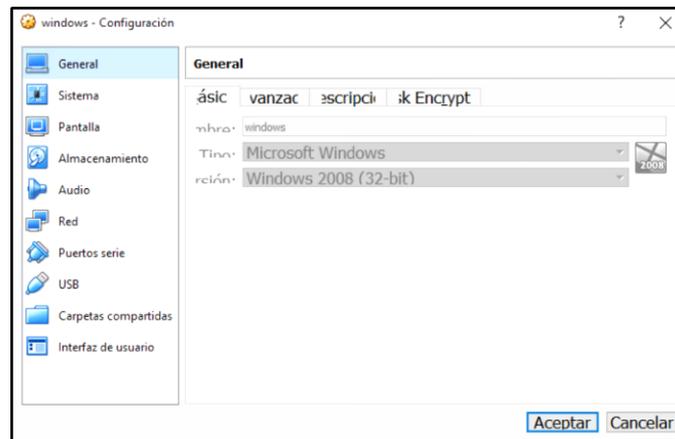


Ilustración 9 – Configuración VirtualBox

La configuración de la máquina virtual es la siguiente:

- 12 Gb de memoria
- 4 procesadores
- 32 Gb espacio en disco
- Adaptador de red puente.

La máquina se llama WIN-FBE5LU6B0OG.

Uno de los principales problemas que he tenido en este apartado ha sido conseguir que mi máquina virtual se pueda conectar a los distintos orígenes de datos con los ETLs. El problema que tenía era que la IP que me daba la máquina virtual era una 192.xxx.xxx.xxx, la cual me dejaba navegar a través del anfitrión sin problemas, pero esa IP no estaba en la red de sistemas que tenía acceso a las bases de datos que necesitaba atacar. En los ajustes de red de mi máquina virtual, intentaba habilitar el adaptador de red para usar un adaptador puente que me diera una IP del tipo 10.xxx.xxx.xxx, que era la que me permitiría acceder a mis orígenes de datos. Pero no me reconocía ningún nombre de adaptador de red. Al investigar un poco más, observé que mi máquina (marca Dell, procesador Intel), no tenía habilitada esa opción. Después de leer documentación, vi que tenía que ir a la BIOS, y en un menú especial de Dell pude habilitar que me dejara instalar adaptadores de red. Tras instalar uno, por fin pude tener en la máquina invitada una IP adecuada.



Ilustración 10 – Configuración Red VirtualBox

2.3.1.2. Instalación de la Base de datos

Como se dijo en puntos anteriores he procedido a instalar un SQL Server 2008 R2 SP2 Express y el SQL Server Management como herramienta cliente de gestión de la base de datos. No es necesario instalar la misma versión en ambas herramientas.

La base de datos se puede descargar de <https://www.microsoft.com/es-es/download/details.aspx?id=30438>

El cliente SQL Server Management Studio 2017 se puede descargar de <https://docs.microsoft.com/es-es/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-2017>

Una vez instaladas ambas versiones hay que asegurarse que los servicios están levantados y funcionando.

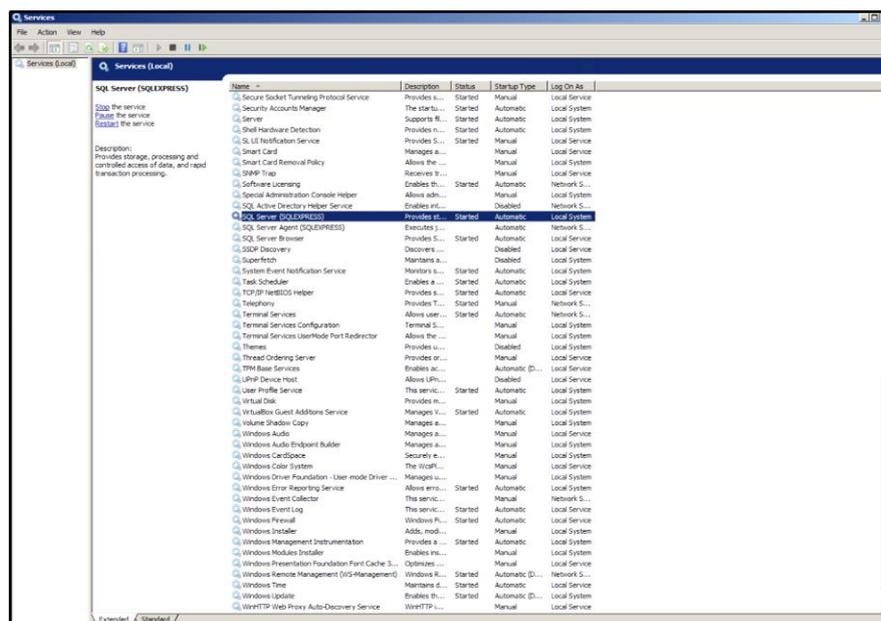


Ilustración 11 – Servicios SQL

El servidor lo he llamado SQLEXPRESS.



Ilustración 12 – Conexión SQL Server

He creado un usuario específico denominado Talend para realizar las conexiones a esta BD, con permisos de lectura/escritura en las tablas y de inicio de sesión.

2.3.1.3. Instalación del software de Extracción, transformación y carga de datos

Como ya he dicho anteriormente, la herramienta elegida es el Talend Open Studio. Ya que es una herramienta muy potente, intuitiva, flexible, gratuita y con una gran comunidad de soporte.

El software de Talend Open Studio Integration puede descargarse en <https://www.talend.com/products/talend-open-studio>

La herramienta no tiene ninguna dificultad para ser instalada mientras se cumplan los requisitos mínimos:

- 3GB mínimo, 4 GB recomendado
- 3GB espacio en disco
- Java 8 JRE Oracle

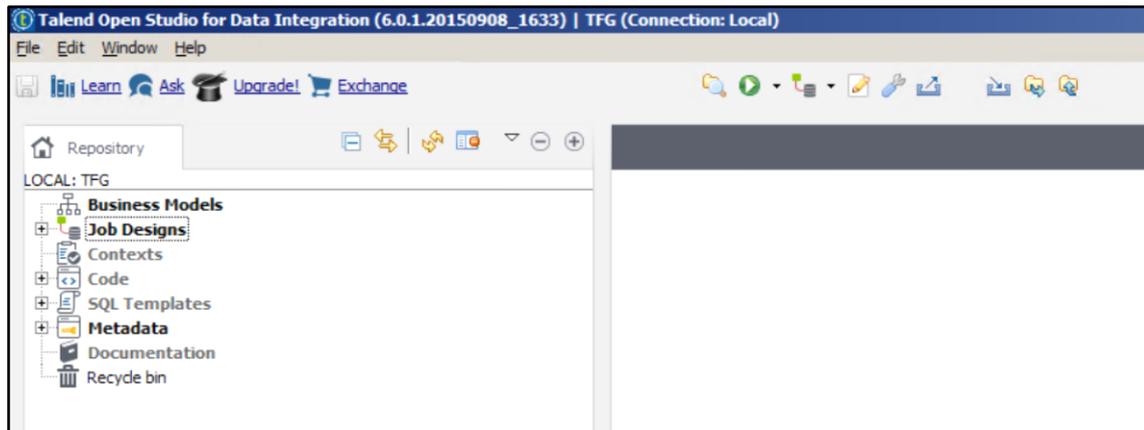


Ilustración 13 – Talend

2.3.1.4. Instalación del software de creación de informes

En este caso, he instalado el Microstrategy Desktop. También ha sido bastante sencillo, basta con descargarlo y seguir un par de pasos de configuración. Se puede descargar de la siguiente url <https://www.microstrategy.com/es/get-started/desktop>.

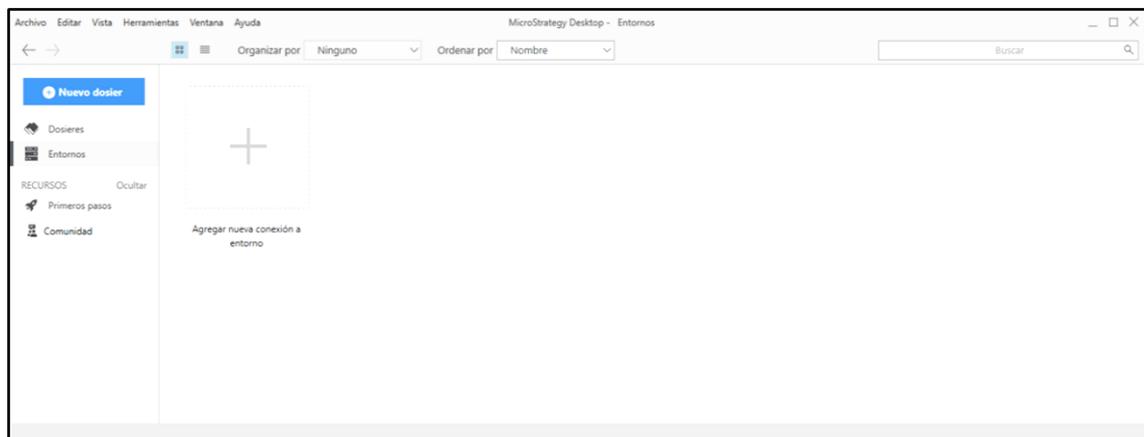


Ilustración 14 – Microstrategy Desktop

2.3.2. Base de datos

En el presente apartado, intentare detallar la secuencia de tareas que hicieron falta para tener una base de datos que cubra las necesidades especificadas anteriormente.

Pasos realizados:

1. Creación de BD

```
USE [master]
GO
CREATE DATABASE [CMI] ON PRIMARY
( NAME = N'CMI', FILENAME = N'c:\Program Files\Microsoft SQL Server\MSSQL10.SQLEXPRESS\MSSQL\DATA\CMI.mdf' , SIZE = 31744KB , MAXSIZE = UNLIMITED, FILEGROWTH = 1024KB )
LOG ON
( NAME = N'CMI_log', FILENAME = N'c:\Program Files\Microsoft SQL Server\MSSQL10.SQLEXPRESS\MSSQL\DATA\CMI_log.ldf' , SIZE = 43264KB , MAXSIZE = 2048GB , FILEGROWTH = 10%)
GO
```

2. Permisos al usuario Talend a esa BD

```
USE [CMI]
GO
CREATE USER [talend] FOR LOGIN [talend] WITH DEFAULT_SCHEMA=[dbo]
GO
```

3. Creación de tablas

4.

4.1. Indicadores

```
USE [CMI]
GO
CREATE TABLE [dbo].[INDICADORES] (
    [Id] [int] IDENTITY(1,1) NOT NULL,
    [Indicador] [nvarchar](500) NULL,
    [Valor] [float] NULL,
    [Objetivo] [float] NULL,
    CONSTRAINT [PK_INDICADORES] PRIMARY KEY CLUSTERED
(
    [Id] ASC
)
) ON [PRIMARY]
GO
```

4.2. Simove

```
USE [CMI]
GO
CREATE TABLE [dbo].[SIMOVE] (
    [travel_id] [nvarchar](20) NULL,
    [travel_date] [date] NULL,
    [vehicle] [nvarchar](10) NULL,
    [driver] [nvarchar](20) NULL,
    [speed] [float] NULL,
    [radius] [float] NULL,
    [accelpos] [float] NULL,
    [accelneg] [float] NULL
) ON [PRIMARY]
GO
```

4.3. Validaciones

```
USE [CMI]
GO

CREATE TABLE [dbo].[VALIDACIONES] (
  [ID] [bigint] NULL,
  [securitydomainid] [bigint] NULL,
  [fecha] [date] NULL,
  [locationx] [float] NULL,
  [locationy] [float] NULL
) ON [PRIMARY]

GO
```

4.4. Ventas

```
USE [CMI]
GO

CREATE TABLE [dbo].[VENTAS] (
  [Id] [int] IDENTITY(1,1) NOT NULL,
  [Fecha] [date] NULL,
  [Precio_Venta] [float] NULL,
  [pass] [nvarchar] (500) NULL,
  CONSTRAINT [PK_VENTAS] PRIMARY KEY CLUSTERED
 (
   [Id] ASC
 )
) ON [PRIMARY]

GO
```

Lanzando todos estos scripts, ya tenemos la base de datos preparada para insertar toda la información que queremos explotar de los distintos orígenes de datos.

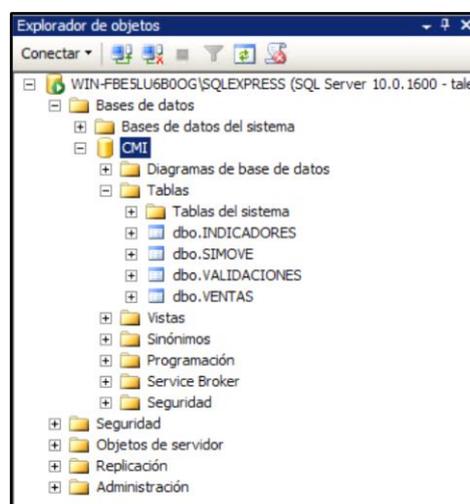


Ilustración 15 – Base de Datos CMI

2.3.3. Creación ETLs

En el presenta apartado, he implementado la solución expuesta en el apartado “2.2.3. Extracción transformación y carga de datos”.

Para ello se han creado 5 trabajos de Talend, los cuatro que atacan a los distintos orígenes y uno adicional llamado Main que ejecuta cada uno de los otros trabajos en el orden deseado.

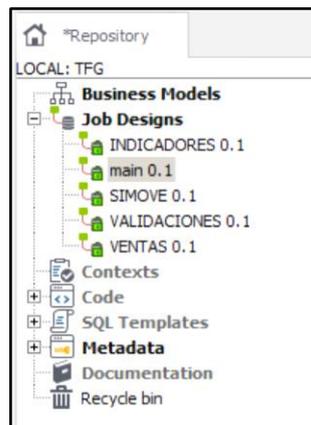


Ilustración 16 – Proyecto Talend TFG



Ilustración 17 – Trabajo Main

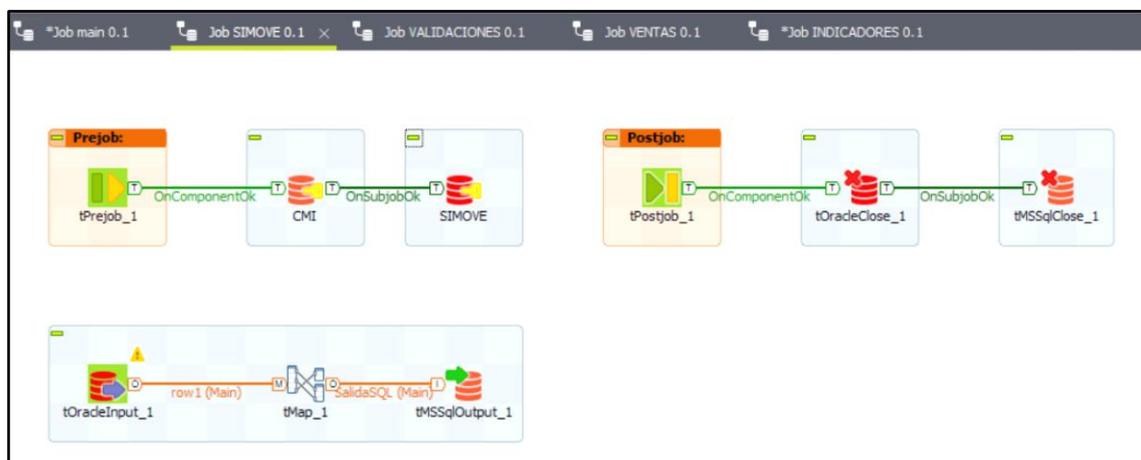


Ilustración 18 – Trabajo SIMOVE

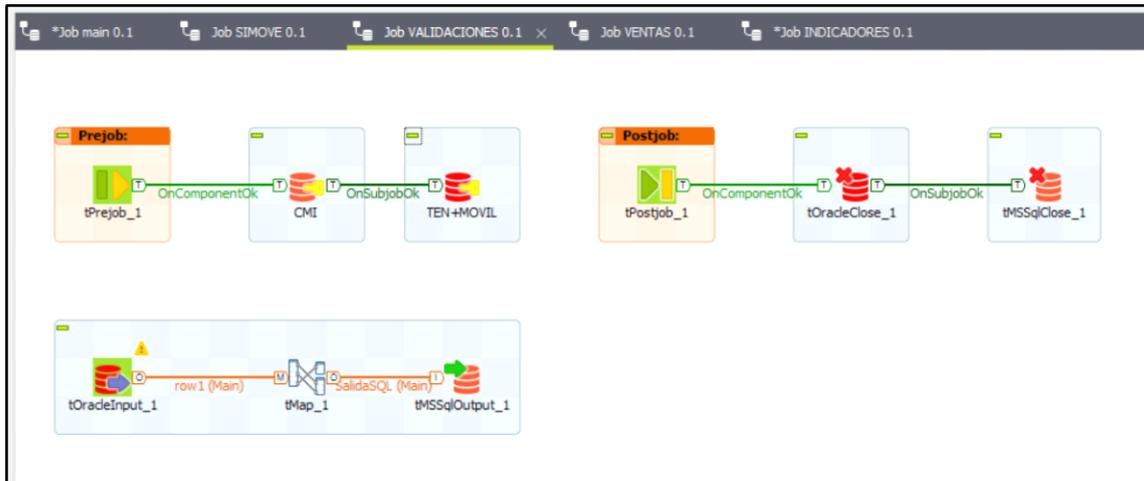


Ilustración 19 – Trabajo Validaciones

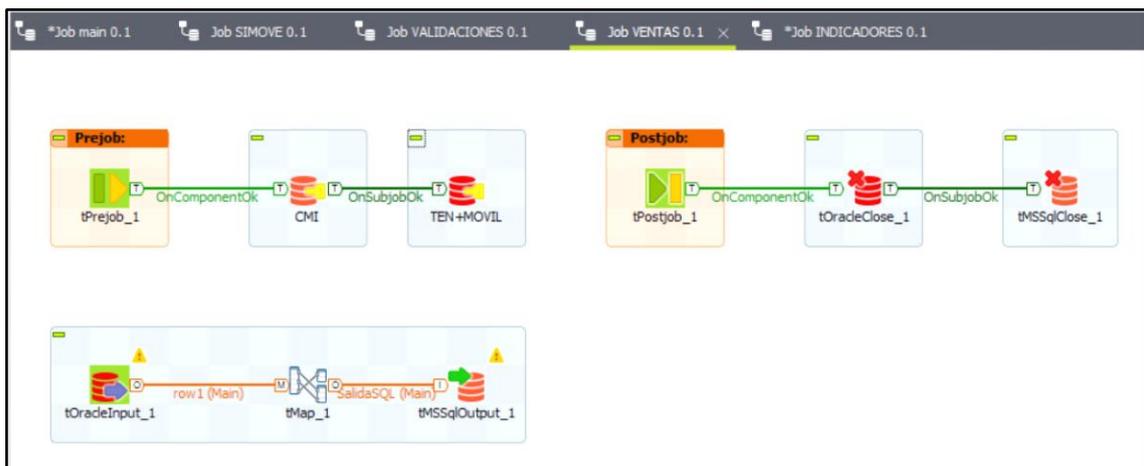


Ilustración 20 – Trabajo Ventas

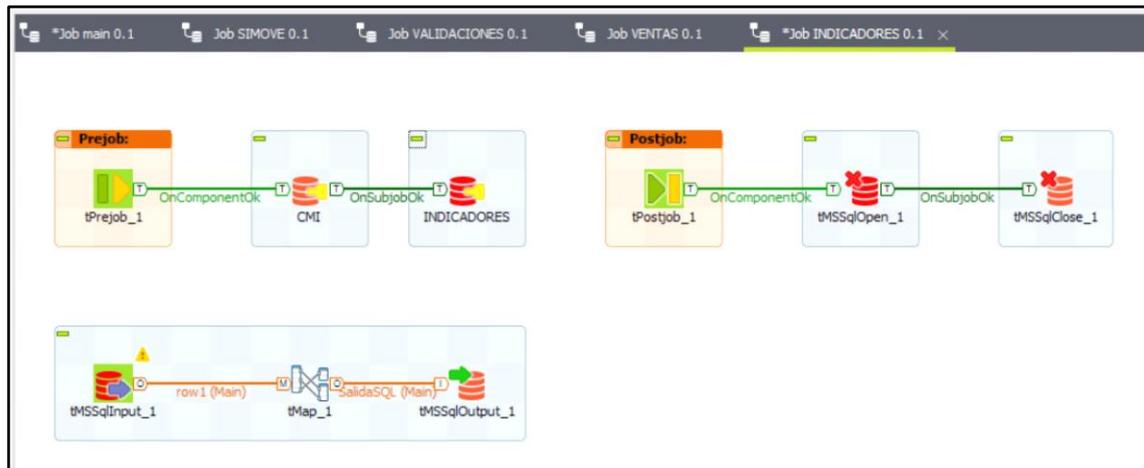


Ilustración 21 – Trabajo Indicadores

A continuación, voy a detallar uno de los trabajos, ya que son bastante similares entre ellos, aunque apunten a orígenes de distinta índole y haya que hacer mapeo de tipos de campos distintos.

Siempre hago un Prejob y un Postjob, la idea es abrir y cerrar las conexiones una sola vez por ejecución, para ser más eficientes. El Prejob abre las conexiones del origen y destino de las BBDD mientras que el PostJob las cierra.

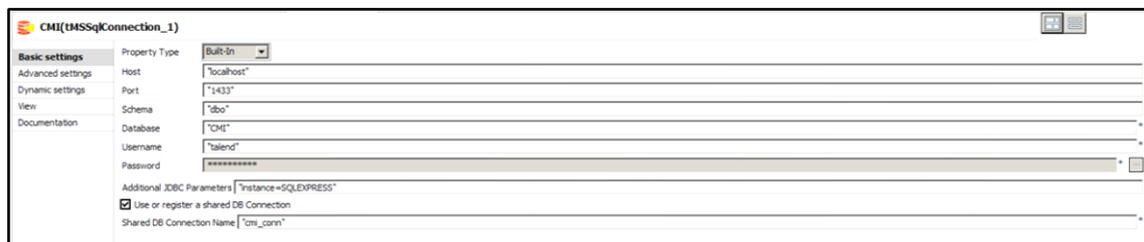


Ilustración 22 – Conexión SQL server CMI

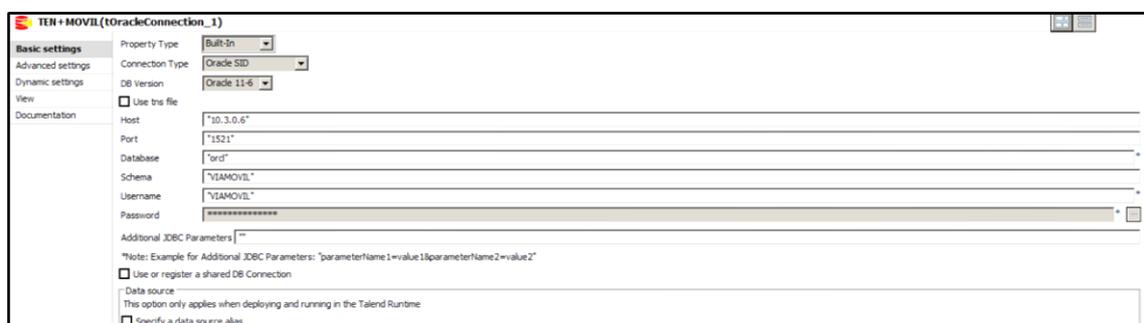


Ilustración 23 – Conexión Oracle VIAMOVIL

La parte principal del trabajo, hay un objeto del tipo Input, ya sea contra Oracle o SQL Server, un objeto que me mapea los tipos de datos, longitudes de cadenas, etc... y un objeto del tipo Output SQL Server para almacenar la información recogida en el Input.

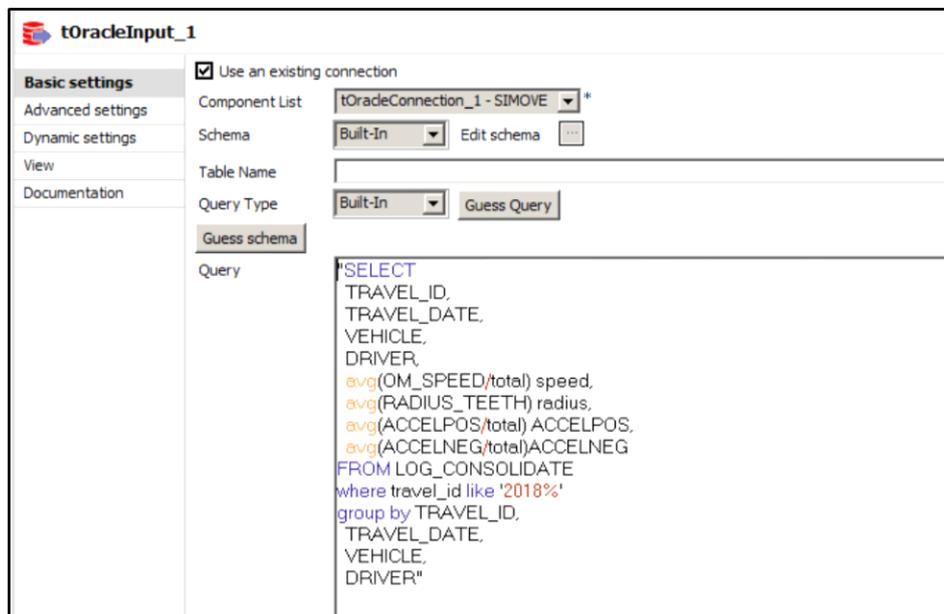


Ilustración 24 – Input Oracle SIMOVE

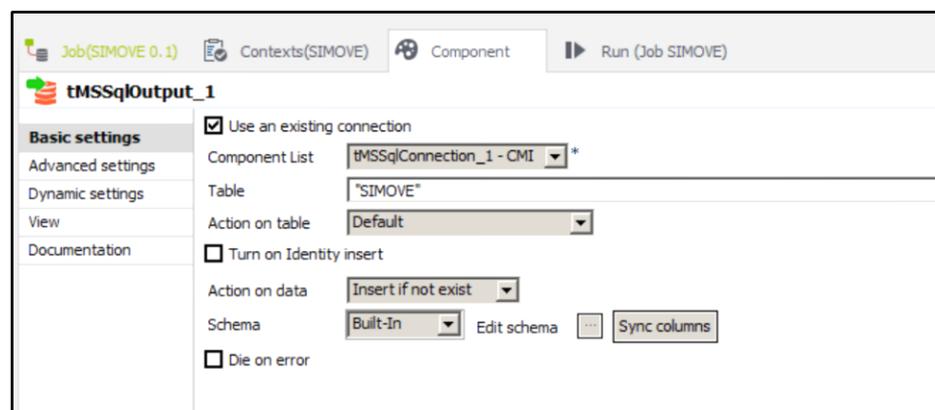


Ilustración 25 – Output SQL Server CMI

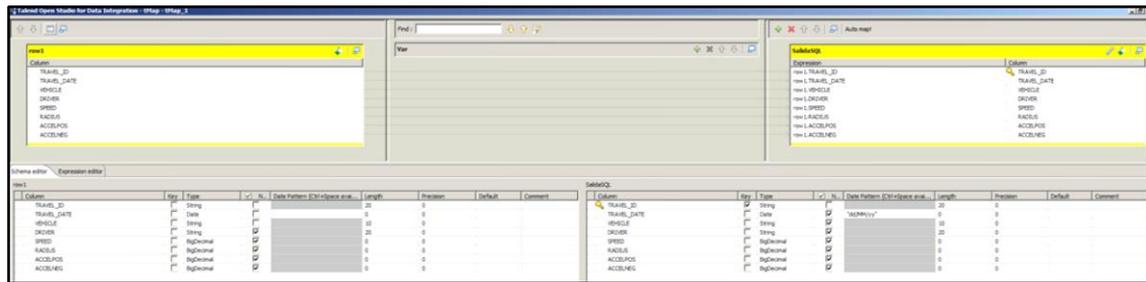


Ilustración 26 – tMap

A continuación, procedo a realizar una carga para comprobar que la implementación es correcta.

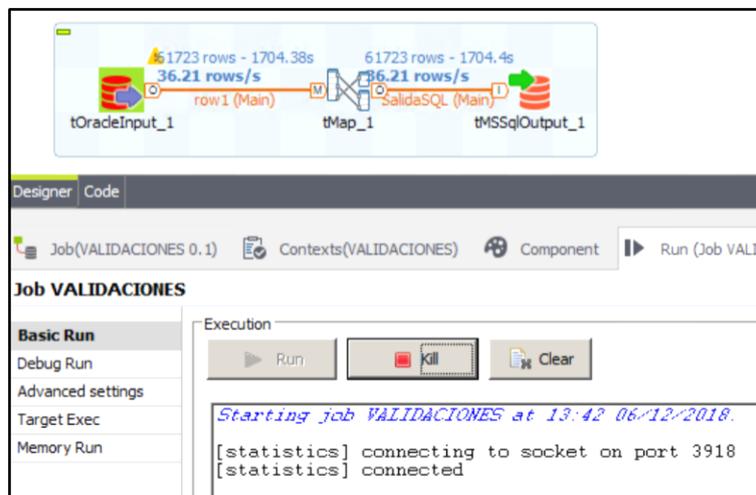


Ilustración 27– Ejecución Job

Para comprobar que toda la carga ha ido de la forma esperada procedo a ejecutar un “Select top 1000 ...” de cada una de las tablas que tenemos en el CMI.

```

SELECT TOP 1000 [Id]
, [Indicador]
, [Valor]
, [Objetivo]
FROM [CMI].[dbo].[INDICADORES]
    
```

	Id	Indicador	Valor	Objetivo
1	1	D1 - Viajes Perdido	99,87	98
2	2	D2 - Viajes Retrasados	99,76	98
3	3	D3 - Viajes Adelantados	99,23	98
4	4	D4 - Viajes Salida Temprana	99,43	98
5	5	D5 - Viajes Salida Tardía	98,7	98
6	6	D6 - Limpieza Vehículos	100	98
7	7	D7 - Reparación Vehículos	100	98
8	8	D8 - Limpieza Paradas	100	98
9	9	D9 - Reparación Paradas	100	98
10	10	D10 - Disponibilidad Expendedoras	97,76	98

Ilustración 28– Datos tabla Indicadores

```

SELECT TOP 1000 [travel_id]
, [travel_date]
, [vehicle]
, [driver]
, [speed]
, [radius]
, [accelpos]
, [accelneg]
FROM [CMI].[dbo].[SIMOVE]
    
```

	travel_id	travel_date	vehicle	driver	speed	radius	accelpos	accelneg
1	2018011809000011811	2018-01-18	118	AC-109	25.5128669598246	0.000711105465	0.111589380407615	-0.111325224365315
2	2018013020000011321	2018-01-30	113	AC-137	24.94460622252064	0.000689115965	0.114127152617046	-0.108344416629688
3	2018021809000012521	2018-02-18	125	AC-388	21.7537082922308	0.000683760055	0.0850255027111756	-0.114560188371567
4	2018030810000010311	2018-03-08	103	AC-245	21.2614519668341	0.000699079254	0.148873136271711	-0.0703032501608775
5	2018032207000011721	2018-03-22	117	AC-206	26.8470643550729	0.000699408958	0.129114071734214	-0.120470698601825
6	2018041214000010211	2018-04-12	102	AC-111	23.4989275856452	0.000707068195	0.148839646647522	-0.135478226095671
7	2018043020000011321	2018-04-30	113	AC-108	26.0521718600356	0.00068561805	0.101348777450306	-0.0919537768194678
8	2018050809000011411	2018-05-08	114	AC-137	23.7119093539391	0.000646290608	0.128792551267391	-0.0882116808284811
9	2018053100000012511	2018-05-31	125	AC-139	23.4399269959012	0.000685915189	0.180326678150575	-0.117723644225752
10	2018060609000011021	2018-06-06	110	AC-114	27.06615660538	0.0006739171	0.114011198763542	-0.124393246953848
11	2018062607000011422	2018-06-26	114	AC-259	22.5919849392614	0.000654264756	0.15275474506155	-0.106505206715893
12	2018072908000011122	2018-07-29	111	AC-238	21.3386175009463	0.000654986737	0.10356506379231	-0.12501267426941
13	2018081014000011611	2018-08-10	116	AC-214	25.2235495589544	0.000691421532	0.165789859749241	-0.114675392543545
14	2018082714000011011	2018-08-27	110	AC-139	22.3682657825512	0.00068013975	0.166795899354237	-0.0955796606811149
15	2018091709000011111	2018-09-17	111	AC-388	25.7737376333982	0.000656876248	0.124556262287535	-0.116022706631172
16	2018100906000011811	2018-10-09	118	AC-69	0	0.000705132999	0	0
17	2018101714000012411	2018-10-17	124	AC-201	23.4938195326667	0.000677893517	0.18570744047348	-0.138832612813239
18	2018010112000012022	2018-01-01	120	AC-212	23.0409689541978	0.000646957076	0.239086424905807	-0.111733126884426
19	2018011822000012621	2018-01-18	126	AC-137	22.1311153381957	0.000650664687	0.113424542852955	-0.0814764843461505
20	2018020713000010912	2018-02-07	109	AC-245	23.4642234102992	0.000681554251	0.132857256996936	-0.157713670232647
21	2018021720000010311	2018-02-17	103	AC-212	22.1134615563502	0.000699063746	0.154363208142112	-0.137823105061926
22	2018030912000011621	2018-03-09	116	AC-124	25.2535270969426	0.000698845665	0.13290668243811	-0.130076190759119
23	2018032416000012521	2018-03-24	125	AC-259	24.5084451646654	0.000682796117	0.134073095805113	-0.0957147949057885
24	2018041012000011512	2018-04-10	115	AC-104	23.6977788362814	0.000651249881	0.181901147572922	-0.0992545404062332
25	2018042713000011221	2018-04-27	112	AC-280	24.3358272348814	0.000709050892	0.146864940403743	-0.118174166296094
26	2018052008000012311	2018-05-20	123	AC-137	26.332701064225	0.00066645575	0.103498054676162	-0.102570166514566
27	2018061109000012411	2018-06-11	124	AC-206	24.7215612226759	0.00069444363	0.152905900088746	-0.109411620565178

Ilustración 29 – Datos tabla Simove

```

SELECT TOP 1000 [ID]
, [securitydomainid]
, [fecha]
, [locationx]
, [locationy]
FROM [CMI].[dbo].[VALIDACIONES]
    
```

	ID	securitydomainid	fecha	locationx	locationy
1	1300247015	1297737680	2018-09-07	NULL	NULL
2	1300948202	569525637	2018-09-07	28,0609727	-16,728489
3	1300506925	373331794	2018-09-07	28,39361	-16,519515
4	1300222747	760477981	2018-09-07	NULL	NULL
5	1300457727	16823351	2018-09-07	NULL	NULL
6	1301785970	204730232	2018-09-08	NULL	NULL
7	1300760291	19544919	2018-09-07	NULL	NULL
8	1305115975	143181113	2018-09-11	NULL	NULL
9	1305494791	114645714	2018-09-11	NULL	NULL
10	1301037097	490157863	2018-09-07	NULL	NULL
11	1302946881	318629776	2018-09-10	NULL	NULL
12	1305387161	762363150	2018-09-11	NULL	NULL
13	1301051513	332928643	2018-09-07	NULL	NULL
14	1303671133	262750040	2018-09-10	NULL	NULL
15	1305690190	53417472	2018-09-11	NULL	NULL
16	1301405565	228395622	2018-09-08	NULL	NULL
17	1301282509	1024498090	2018-09-08	28,3569724	-16,3693923
18	1302246325	1295114493	2018-09-09	NULL	NULL
19	1301193138	965504522	2018-09-08	NULL	NULL
20	1301108757	134205	2018-09-07	NULL	NULL
21	1301726218	383802807	2018-09-08	28,4140221	-16,5530368
22	1302884242	372893017	2018-09-10	NULL	NULL
23	1302391760	546910443	2018-09-09	NULL	NULL
24	1306013460	53044581	2018-09-11	NULL	NULL
25	1302443806	117316	2018-09-09	NULL	NULL
26	1303911276	221290918	2018-09-10	NULL	NULL
27	1302789008	73801803	2018-09-10	28,3695545	-16,7140962
28	1305740089	13769872	2018-09-11	NULL	NULL
29	1301549096	41182583	2018-09-08	NULL	NULL
30	1304088685	894192009	2018-09-10	28,4524334	-16,3004336
31	1303115007	909852556	2018-09-10	NULL	NULL
32	1304539473	789821718	2018-09-11	NULL	NULL
33	1303304747	32420835	2018-09-10	NULL	NULL

Ilustración 30 – Datos tabla Validaciones

```

SELECT TOP 1000 [Id]
, [Fecha]
, [Precio_Venta]
, [pass]
FROM [CMI].[dbo].[VENTAS]
    
```

	Id	Fecha	Precio_Venta	pass
1	1	2018-10-05	38	Abono Mensual
2	2	2018-10-05	1,1	Sencillo Tranvía
3	3	2018-10-02	10	Abono Joven Semanal
4	4	2018-10-02	38	Abono Mensual
5	5	2018-10-02	30	Abono Joven Mes
6	6	2018-10-02	15	BonoVia 15
7	7	2018-10-02	15	BonoVia 15
8	8	2018-10-03	10	Abono Joven Semanal
9	9	2018-10-03	15	BonoVia 15
10	10	2018-10-05	15	BonoVia 15
11	11	2018-10-05	38	Abono Mensual
12	12	2018-10-05	15	BonoVia 15
13	13	2018-10-05	15	Bono E. Universitario
14	14	2018-10-05	25	BonoVia 25
15	15	2018-10-05	15	Bono E. Universitario
16	16	2018-10-05	47	Bono Residente Canario
17	17	2018-10-05	5	Bono 5 viajes
18	18	2018-10-05	1,1	Sencillo Tranvía
19	19	2018-10-02	1,1	Sencillo Tranvía
20	20	2018-10-02	1,1	Sencillo Tranvía
21	21	2018-10-02	1,1	Sencillo Tranvía
22	22	2018-10-04	15	BonoVia 15
23	23	2018-10-04	15	BonoVia 15

Ilustración 31 – Datos tabla Ventas

2.3.4. Programación ETLs

En el presente apartado explicaré como he automatizado las cargas de los ETLs para tener datos actualizados de forma diaria por las necesidades del proyecto.

Para ello, lo primero que he realizado es una exportación del trabajo “main” del talend, el cual como ya he comentado anteriormente se encarga de lanzar todos los trabajos de carga con un orden lógico.

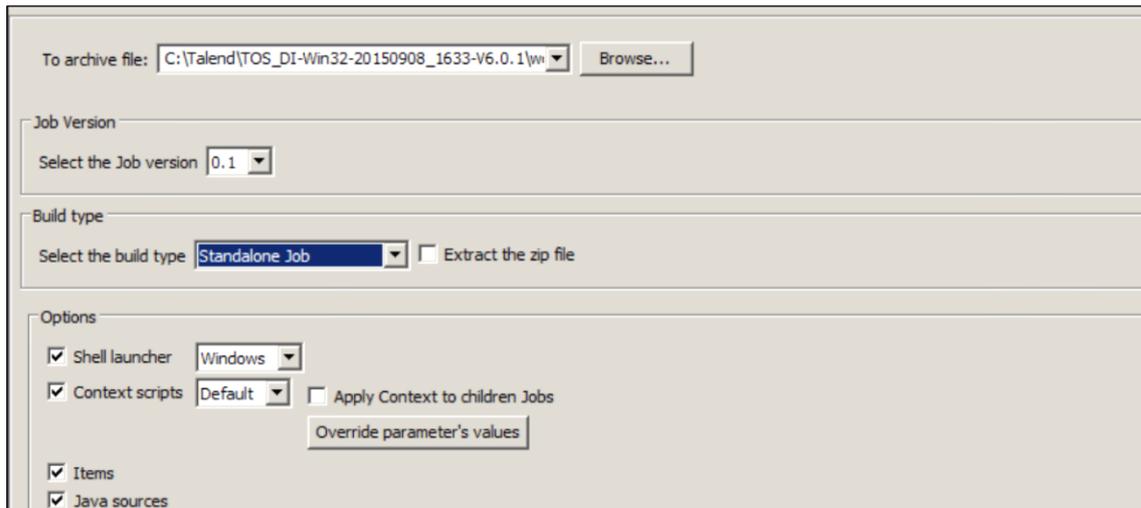


Ilustración 32 – Exportar trabajo Talend

Una vez exportado el trabajo, he creado un pequeño script en un fichero de extensión “.bat”, que ejecutará el trabajo.

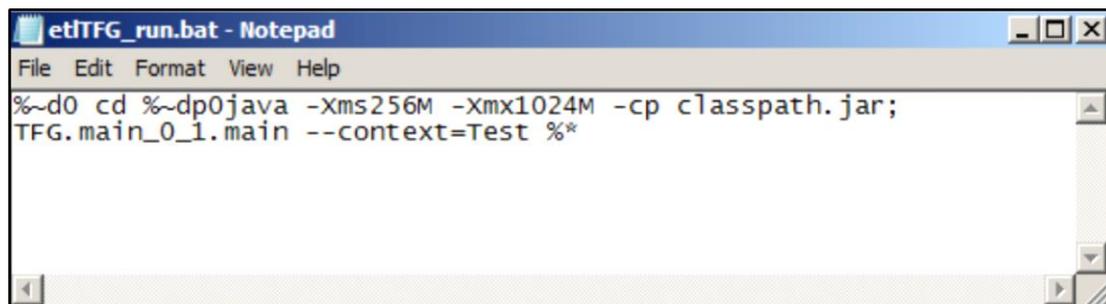


Ilustración 33 – Script trabajo Talend

Para realizar la planificación del trabajo, he barajado dos opciones como son las tareas del motor de base de datos de SQL Server y las tareas de Windows. Finalmente, me he decantado por las tareas de Windows por tener un cierto control en el visualizador de eventos, aunque con los trabajo de SQL podría ver el historial de ejecuciones, ambas opciones eran igual de buenas.

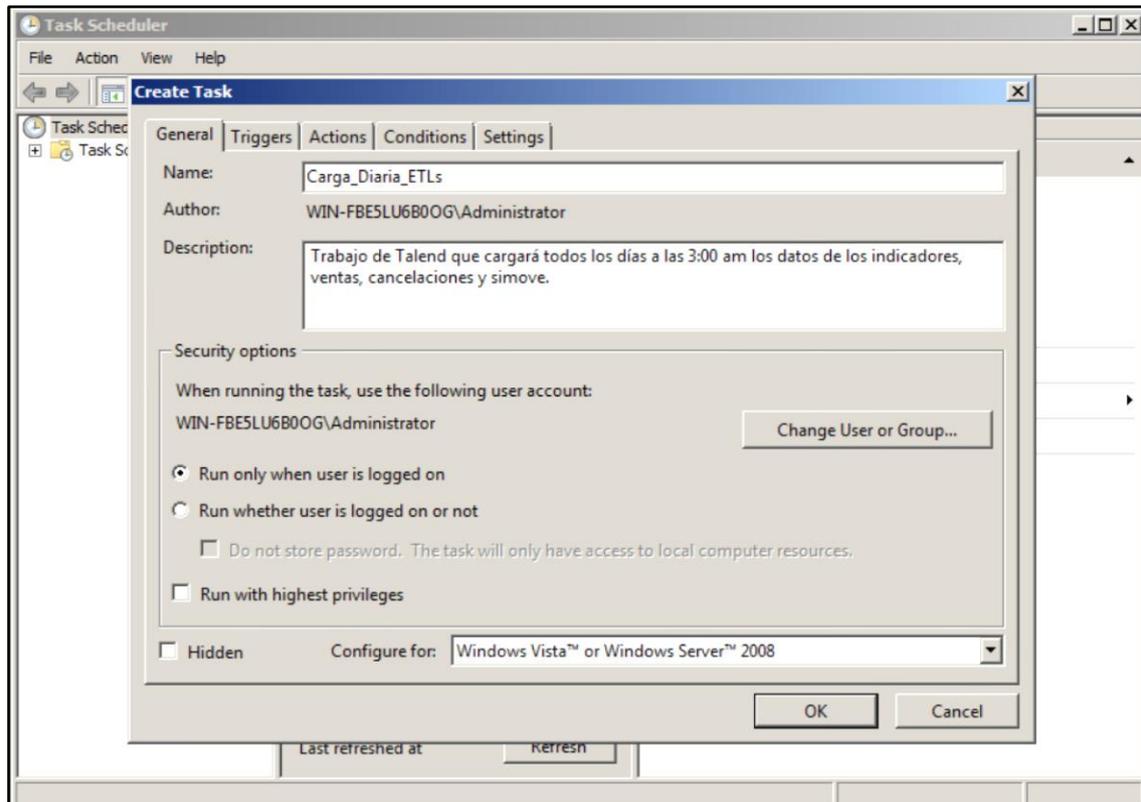


Ilustración 34 – Creación de tarea parte 1

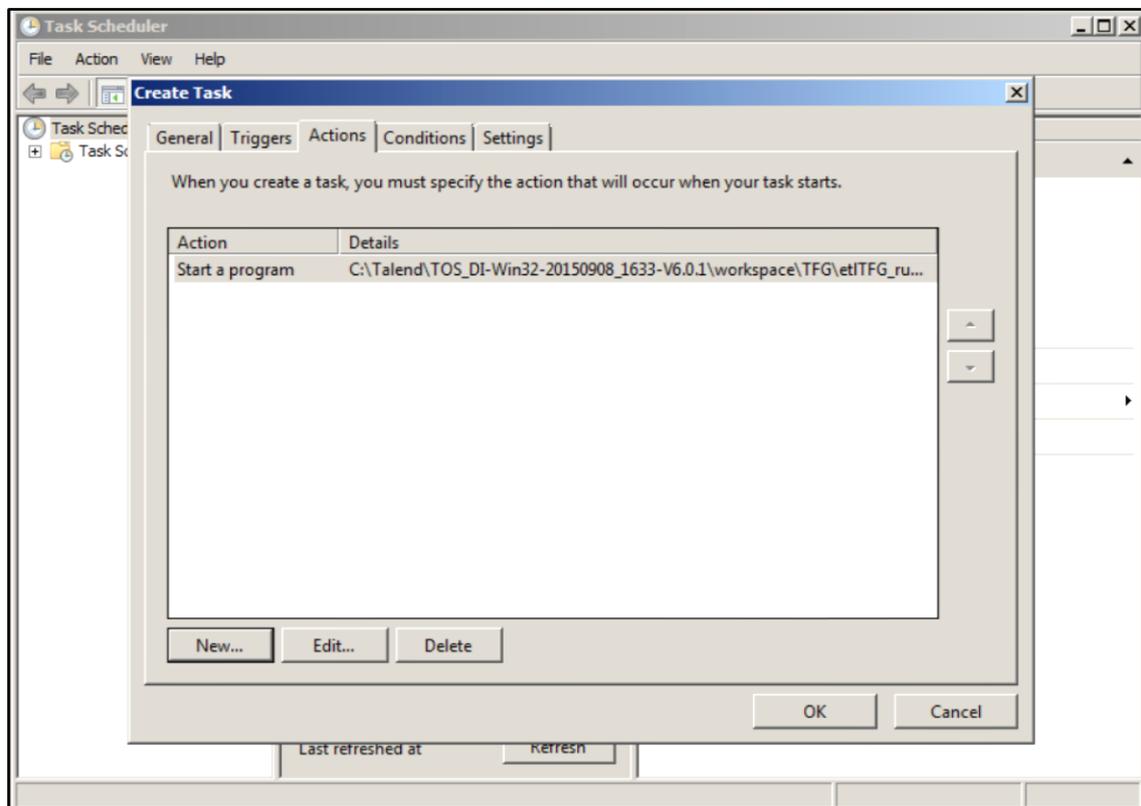


Ilustración 35 – Creación de tarea parte 2

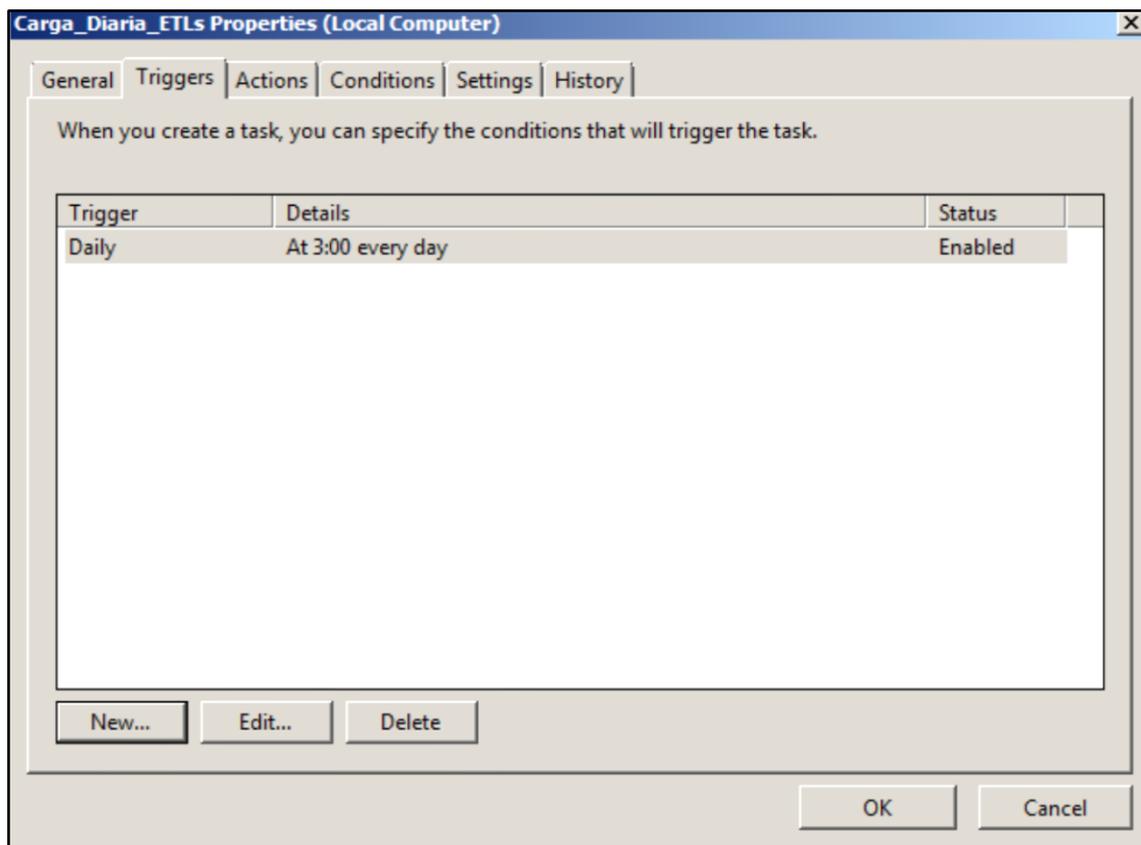


Ilustración 36 – Creación de tarea parte 3

2.3.5. Creación CMI

En el presente apartado voy a mostrar los resultados en forma de un cuadro de mando integral, donde se podrá ver la última capa del proyecto, en el que el usuario final puede analizar de una forma visual los resultados de los indicadores de la empresa.

Para ello, como ya se ha mencionado anteriormente, usaré la herramienta Microsoft Desktop. Creando lo que ellos llaman Dosieres, que no es otra cosa que un dashboard o cuadro de mando.



Ilustración 37 – Dossier CMI

Una vez creado el dossier, cree los distintos conjuntos de datos contra mi base de datos CMI.

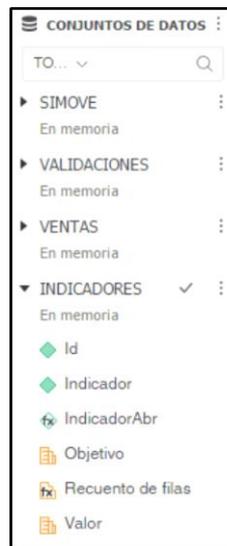


Ilustración 38 – Conjuntos de datos

Cada uno de estos conjuntos de datos, cargaba la información de la tabla correspondiente en el CMI.

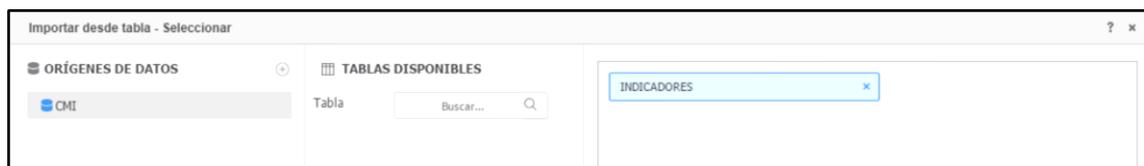


Ilustración 39 – Ejemplo de conjunto de datos

A continuación, únicamente quedaba darle forma según lo diseñado anteriormente con la herramienta Balsamiq Mockups.



Ilustración 40 – Informe de indicadores

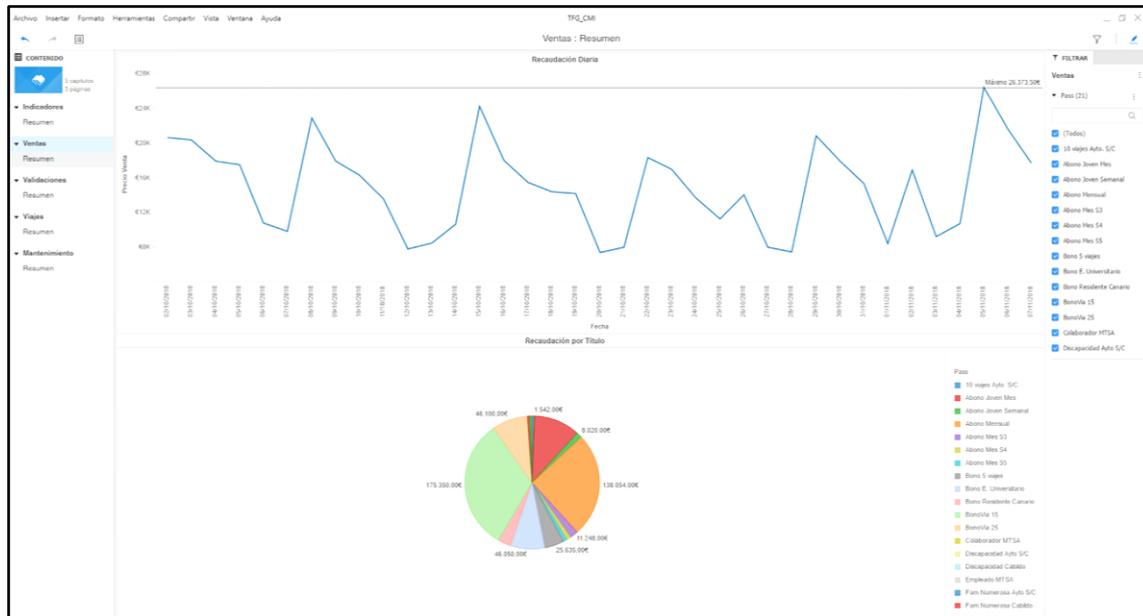


Ilustración 41 – Informe de ventas Diarias



Ilustración 42 – Informe de viajes diarios

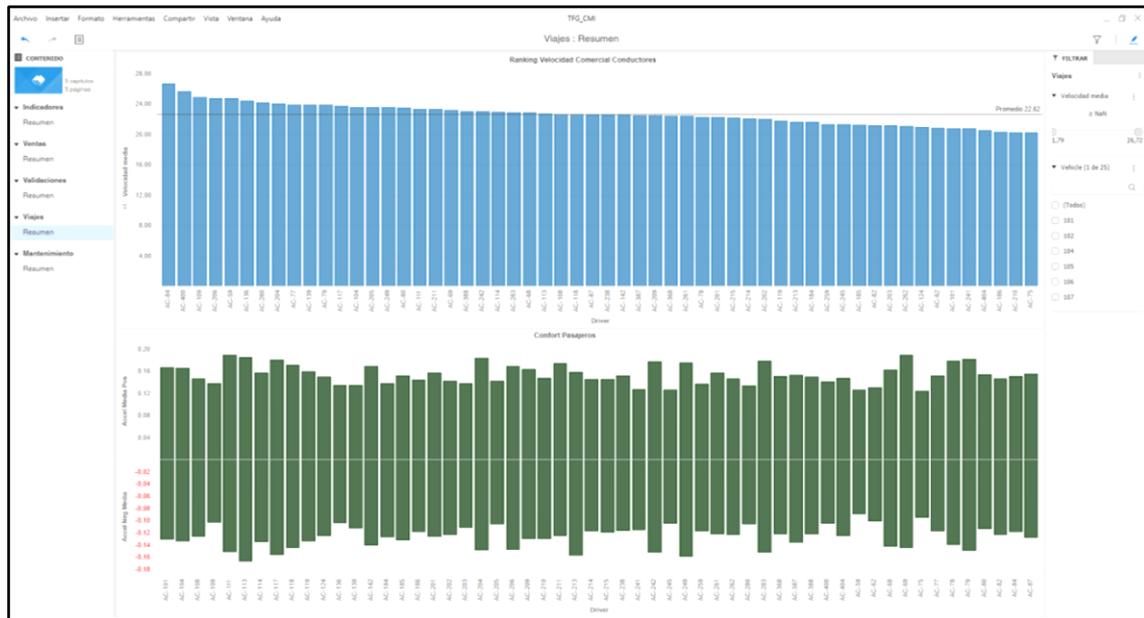


Ilustración 43 – Informe de velocidad por conductor

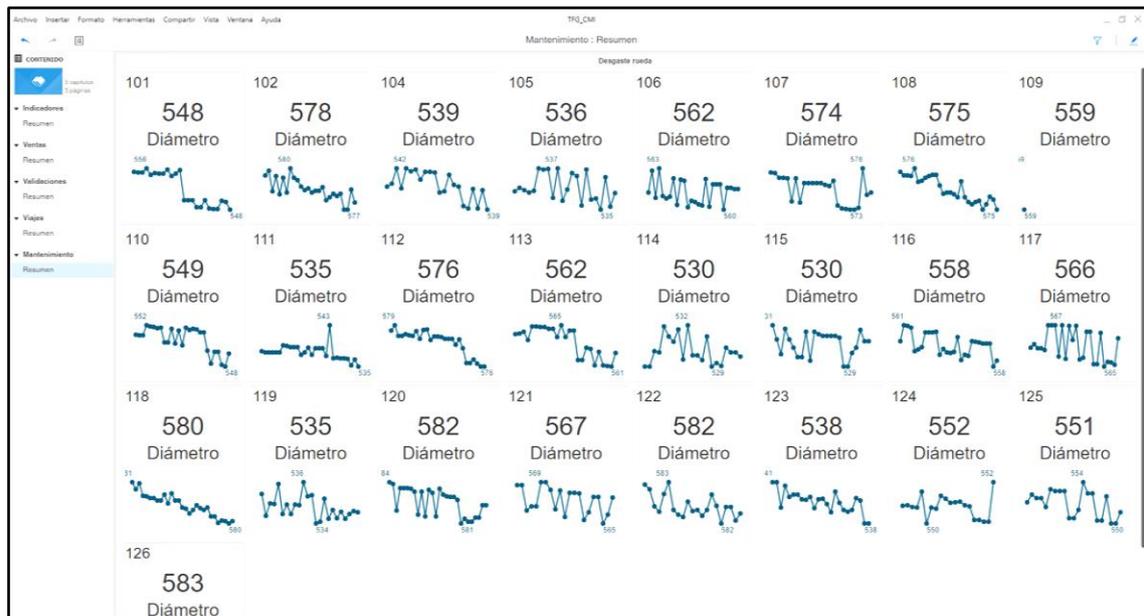


Ilustración 44 – Informe de desgaste de las ruedas de los vehículos

3. Conclusiones

La primera conclusión es que se ha conseguido el objetivo principal del TFG, que es la aplicación y puesta en práctica de todos los conocimientos adquiridos durante los cuatro años del Grado de Ingeniería Informática.

Este TFG, ha usado como guía para su desarrollo el ciclo de vida del software: requisitos, análisis, diseño, implementación, pruebas y validación, lo que ha supuesto un gran esfuerzo, ya que no es únicamente una implementación sino el desarrollo completo de un producto que perfectamente podría ser real.

Personalmente, creo que la experiencia ha sido enriquecedora, exigente por los plazos, pero satisfactoria con el resultado final. He tenido que aprender nuevas tecnologías, leer documentación y sobre todo buscar soluciones tecnológicas que satisfagan las necesidades de la fase de requisitos.

Durante el desarrollo del TFG en coordinación con el tutor, he realizado un plan de proyecto, que me ayudó muchísimo a alcanzar los objetivos y planificarme en tiempo y alcance. Me he quedado muy satisfecho que no tener que modificar ni los objetivos, ni los tiempos establecidos en el plan de proyecto durante la realización del mismo, lo que indica que fueron bien dimensionados. Consiguiendo realizar todos los objetivos con éxito.

Creo que el reto mayor con el que me enfrenté durante el TFG, fue la curva de aprendizaje, al usar tecnologías nuevas para mí, Talend y Microstrategy. El Talend me ha sorprendido gratamente por su potencia, versatilidad y facilidad. Estoy encantado de saber usarlas ahora.

4. Líneas de futuro

Creo que el TFG es el comienzo de un sistema de información que permita a los mandos intermedios y superiores ver de una forma rápida e inequívoca, el estado de toda la empresa, y en caso de ser necesario, poder atomizar la información para su posterior análisis.

Para líneas futuras, habría que integrar nuevos sistemas como el SAE, GMAO, Heures, etc... que permitan nuevos indicadores.

También, habría que añadir orígenes de datos externos al sistema, como podrían ser meteorológicos, redes sociales, etc... con el objetivo de crear un sistema predictivo que nos permita, mediante la modificación de parámetros, como podría ser el número necesario de tranvías y conductores necesarios para prestar un servicio, etc... ver el comportamiento de los indicadores.

Además, creo que el sistema sería fácilmente exportable a otros operadores ferroviarios, ya que la mayoría usan las mismas herramientas y necesitan medir indicadores similares.

5. Glosario

Término	Descripción
MTSA	Metropolitano de Tenerife SA
IeIS	Informática e Ingeniería de Sistema
BI	Bussiness Intelligence
CIO	Chief Information Officer
TFG	Trabajo Fin de Grado
CMI	Cuadro de Mando Integral
ETL	Extract, Transformation and Load
FU	Frenada Urgencia
IHM	Interfaz Hombre Máquina
ECIT	Excelentísimo Cabildo Insular de Tenerife
GMAO	Gestión de Mantenimiento Asistido por Ordenador
Heures	Sistema de planificación horaria
Validaciones	Cada uno de los viajes realizados por los usuarios

Tabla 18 – Glosario

6. Bibliografía

La bibliografía que se ha usado, ha sido para adquirir el conocimiento necesario de las distintas herramientas, en ninguna parte del TFG se hace mención directa a textos que no son propios del autor de la TFG.

- WEB:
https://info.talend.com/rs/talend/images/CB_EN_DI_Cookbook_DataIntegration.pdf?_ga=2.106179851.904179080.1545899012-1886391384.1538724110 - 06/11/2018
- WEB:
<http://www2.microstrategy.com/producthelp/10.4/manuals/en/AnalyticsDesktopUserGuide.pdf> - 17/12/2018