



UNIVERSIDAD CÉSAR VALLEJO

FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

**Desarrollo de Inteligencia de Negocios con un modelo de
Machine Learning para la gestión de la flota vehicular**

TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE INGENIERO DE
SISTEMAS

AUTOR:

Vergara Pinedo, Juan Luis (ORCID: 0000-0002-9233-4029)

ASESOR:

Dr. Hilario Falcón, Francisco Manuel (ORCID: 0000-0003-3153-9343)

LÍNEA DE INVESTIGACIÓN:

Sistemas de Información y Comunicaciones

LIMA – PERÚ

2021

DEDICATORIA

La presente tesis es dedicada a mi familia por darme una formación basada en valores, su apoyo en los momentos con mayor dificultad a lo largo de mi vida y por guiarme en mi formación profesional.

AGRADECIMIENTOS

A mi padre por sus consejos y por enseñarme a ser perseverante para cumplir mis metas y a mi madre por su gran amor y por formarme con los valores esenciales para ser una persona correcta. Y a ambos les agradezco por nunca darse por vencidos.

Índice de Contenidos

Contenido

I.	INTRODUCCIÓN.....	1
II.	MARCO TEÓRICO.....	4
III.	METODOLOGÍA.....	39
3.1	Tipo y diseño de la investigación.....	40
3.2	Variables y operacionalización.....	40
3.3	Población, muestra y muestreo.....	43
3.4	Técnicas e instrumentos de recolección de datos, validez y confiabilidad.....	44
3.5	Métodos de análisis de datos.....	45
3.6	Aspectos éticos.....	46
IV.	RESULTADOS.....	47
V.	DISCUSIÓN.....	75
VI.	CONCLUSIONES.....	79
VII.	RECOMENDACIONES.....	82
	REFERENCIAS.....	84
	ANEXOS.....	90

Índice de tablas

Tabla N°1 Matriz de operacionalización de las variables de la investigación	41
Tabla N°2 Indicadores	42
Tabla N°3 Análisis descriptivo del indicador porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	48
Tabla N°4 Análisis descriptivo del indicador porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	49
Tabla N°5 Análisis descriptivo del indicador porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	50
Tabla N°6 Análisis descriptivo del indicador porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	51
Tabla N°7 Análisis descriptivo del indicador porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	52
Tabla N°8 Análisis descriptivo del indicador rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	53
Tabla N°9 Test de Shapiro Wilk de porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	55
Tabla N°10 Test de Shapiro Wilk de porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	57
Tabla N°11 Test de Shapiro Wilk de porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	58
Tabla N°12 Test de Shapiro Wilk de porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	60

Tabla N°13 Test de Shapiro Wilk de porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	62
Tabla N°14 Test de Shapiro Wilk de rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	63
Tabla N°15 Test de Wilcoxon para el porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	66
Tabla N°16 Test de Wilcoxon para el porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	67
Tabla N°17 Test de Wilcoxon para el porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	69
Tabla N°18 Test de Wilcoxon para el porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	70
Tabla N°19 Test de Wilcoxon para el porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	72
Tabla N°20 Test de Wilcoxon para el rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	74

Índice de figuras

Figura N°1 Arquitectura Data Mart	13
Figura N°2 Arquitectura Data Warehouse	14
Figura N°3 Star Schema & OLAP	16
Figura N°4 Tipo de gráficos para visualización de datos	17
Figura N°5 Regresión Lineal	22
Figura N°6 Regresión Logística	22
Figura N°7 Decision Tree	23
Figura N°8 Support Vector Machine	24
Figura N°9 K Near Neighbors	24
Figura N°10 K-Means	25
Figura N°11 Coeficiente alfa	44
Figura N°12 Media de porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning	49
Figura N°13 Media de porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning	50
Figura N°14 Media de porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning	51
Figura N°15 Media de porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning	52
Figura N°16 Media de porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning	53
Figura N°17 Media de rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning	54
Figura N°18 Test de Shapiro Wilk de porcentaje del costo por consumo de combustible antes del desarrollo de la solución de BI usando un modelo de machine learning.	56

Figura N°19 Test de Shapiro Wilk de porcentaje del costo por consumo de combustible después del desarrollo de la solución de BI usando un modelo de machine learning.	56
Figura N°20 Test de Shapiro Wilk de porcentaje del costo por infracción antes del desarrollo de la solución de BI usando un modelo de machine learning.	57
Figura N°21 Test de Shapiro Wilk de porcentaje del costo por infracción después del desarrollo de la solución de BI usando un modelo de machine learning.	58
Figura N°22 Test de Shapiro Wilk de porcentaje del costo por lavado antes del desarrollo de la solución de BI usando un modelo de machine learning.	59
Figura N°23 Test de Shapiro Wilk de porcentaje del costo por lavado después del desarrollo de la solución de BI usando un modelo de machine learning.	59
Figura N°24 Test de Shapiro Wilk de porcentaje del costo por mantenimiento antes del desarrollo de la solución de BI usando un modelo de machine learning.	61
Figura N°25 Test de Shapiro Wilk de porcentaje del costo por mantenimiento después del desarrollo de la solución de BI usando un modelo de machine learning.	61
Figura N°26 Test de Shapiro Wilk de porcentaje del costo por parchado de llantas antes del desarrollo de la solución de BI usando un modelo de machine learning.	62
Figura N°27 Test de Shapiro Wilk de porcentaje del costo por parchado de llantas después del desarrollo de la solución de BI usando un modelo de machine learning.	63
Figura N°28 Test de Shapiro Wilk de rendimiento antes del desarrollo de la solución de BI usando un modelo de machine learning.	64
Figura N°29 Test de Shapiro Wilk de rendimiento después del desarrollo de la solución de BI usando un modelo de machine learning.	64
Figura N°30 Test de Wilcoxon para el porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning.	66

- Figura N°31 Test de Wilcoxon para el porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning. 68
- Figura N°32 Test de Wilcoxon para el porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning. 69
- Figura N°33 Test de Wilcoxon para el porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning. 71
- Figura N°34 Test de Wilcoxon para el porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning. 72
- Figura N°35 Test de Wilcoxon para el rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning. 74

RESUMEN

La presente investigación se llevó a cabo con el fin de comprobar el resultado del desarrollo de inteligencia de negocios con un modelo de machine learning para la gestión de la flota vehicular, el cual ha sido realizado en base a la información de la operatividad de la flota vehicular.

Esta investigación fue de tipo aplicada, mediante un diseño experimental, específicamente del tipo cuasi experimental. La población utilizada, estuvo compuesta por la información de la operatividad de la flota vehicular recopilada mes a mes. La recopilación de datos fue realizada a través de la ficha de observación.

Los resultados de la presente investigación confirman que el desarrollo de inteligencia de negocios con un modelo de machine learning tuvo un efecto realmente positivo para la gestión de la flota vehicular; en cuanto al porcentaje del costo por consumo de combustible antes fue de 0.93392 y después se redujo a 0.91644, el porcentaje del costo por infracción antes fue 0.0023514 y después se redujo a 0.002196, el porcentaje del costo por lavado antes fue 0.014362 y después se redujo a 0.014250, el porcentaje del costo por mantenimiento antes fue 0.97023 y después se redujo a 0.94063, el porcentaje del costo por parchado de llantas antes fue 0.0011252 y después se redujo a 0.0010928, el rendimiento antes fue 37.2 y después aumento a 37.39.

Palabras clave: Inteligencia de negocios, machine learning, flota vehicular, modelos de predicción, reducción de costos.

ABSTRACT

This research was carried out in order to verify the result of the development of business intelligence with a machine learning model for the management of vehicle fleet, which has been carried out based on information on the operation of the vehicle fleet.

This research was of an applied type, through an experimental design, specifically of the quasi-experimental type. The population used was made up of the information on the operation of the vehicle fleet collected month by month. The data collection was carried out through the observation file.

The results of the current research confirm that the development of business intelligence with a machine learning model had a really positive effect on the management of the vehicle fleet; Regarding the percentage of the cost per fuel consumption before it was 0.93392 and then it was reduced to 0.91644, the percentage of the cost per violation before was 0.0023514 and then it was reduced to 0.002196, the percentage of the cost per wash before was 0.014362 and then it was reduced at 0.014250, the maintenance cost percentage before was 0.97023 and then it was reduced to 0.94063, the tire patch cost percentage before was 0.0011252 and after it was reduced to 0.0010928, the performance before was 37.2 and then increased to 37.39.

Key words: Business intelligence, machine learning, vehicle fleet, prediction models, cost reduction.

I. INTRODUCCIÓN

La empresa cuenta su propia flota vehicular a nivel de Lima y Callao, la cual es gestionada por la División de Gestión de Servicios. Se cuenta con 180 vehículos, los cuales están asignados a las distintas unidades organizacionales, y cada una de estas tiene como obligación brindar información de los costos relacionados a la utilización de sus vehículos asignados, la División de Gestión de Servicios toma en cuenta el costo variable de operación y el nivel de servicio de la flota vehicular para su respectiva gestión, en base a la operatividad de los vehículos. Actualmente la División de Gestión de Servicios cuenta con una herramienta que permite centralizar la información de la flota vehicular, esta herramienta permite almacenar los datos de cada proceso operativo de la gestión de la flota vehicular, pero esta herramienta solo genera reportes de los datos existentes, por lo que la herramienta solo almacena la información y no brinda apoyo para la gestión de la flota vehicular de manera efectiva. Cuando en la División de Gestión de Servicios se requiere algún informe de uno, varios o todos los procesos, se generan múltiples reportes de cada proceso, por lo que se producen dispersas fuentes de información para ser analizadas. Entonces se extraen, consolidan, gestionan e interpretan los datos de forma manual, provocando una mala gestión de la información, ya que así no se tiene un aceptable conocimiento de la situación actual o pasada de los procesos de la gestión de la flota vehicular. Así mismo se realizan informes estadísticos solo con las funcionalidades que provee la herramienta Microsoft Excel, estos informes poseen formatos informales. Todo lo mencionado anteriormente perjudica e impide conseguir una óptima mejora continua en la planificación de estrategias y objetivos, basados en el costo variable de operación y el nivel de servicio. A su vez todos estos factores pueden ser utilizados para pronosticar cuando sería conveniente que un vehículo se diera de baja en un determinado lapso de tiempo, en base a la información gestionada. Se ha identificado que la División de Gestión de Servicios no posee un sólido conocimiento de la información de los procesos de la gestión de la flota vehicular, afectando el proceso de toma de decisiones y generando posibles incrementos en el costo variable de operación y decremento en el nivel de servicio, además cabe precisar que este problema es de suma importancia, ya que esto afecta negativamente a los costos de la flota de vehículos y a la actividad financiera del estado.

Entre las múltiples causas que pueden estar originando este problema, las que se han detectado es que se gestiona la información de la flota vehicular solo con apoyo de herramientas informáticas como Microsoft Excel, Outlook y SharePoint. La herramienta para la gestión de la flota vehicular solo genera reportes de los datos en archivos Excel, estos datos y otros de distintas fuentes son extraídos, consolidados, gestionados e interpretados de forma manual, lo cual requiere tiempo y esfuerzo. Entonces se tiene una gran cantidad de datos que no son utilizados de manera oportuna. No se tiene una herramienta que permita gestionar la información de los distintos procesos de la gestión de la flota de manera eficiente, es decir no se tiene una herramienta que permita gestionar los datos correctos y necesarios, que deben ser analizados a nivel histórico, no se cuenta con reportes de informes estadísticos y comparativos, tampoco existe una correcta visualización de los datos para conocer y comprender el costo variable de operación y el nivel de servicio. No hay un correcto análisis de la información de gestión de la flota vehicular en la División de Gestión de Servicios.

Las consecuencias que puede traer este problema son muchas, entre las que se han detectado están, la extracción de distintas fuentes de datos genera pérdida de tiempo y esfuerzo, no se tiene conocimiento sobre los costos variables de operación y del nivel de servicio, se desconoce la situación actual de cada proceso de la gestión de flota y la evolución operativa de cada uno de estos, no se pueden planificar estrategias eficientes que permitan ayudar a optimizar cada proceso y tampoco tomar medidas operativas para reducir los costos variables de operación e incrementar el nivel de servicio, y no se puede hacer un pronóstico adecuado de cuando un vehículo se debe dar de baja. Por lo tanto, existe dificultad para que la División de Gestión de Servicios pueda realizar una correcta gestión de la flota vehicular.

Lo que se puede hacer es procesar toda la información de la gestión de la flota vehicular, utilizando Business intelligence con un modelo de Machine Learning, para que obtener un soporte que permita obtener conocimiento del nivel de servicio y de los costos variables de operación, en base a confiables fuentes de información y así la División de Gestión de Servicios pueda tomar decisiones que generen ventajas competitivas con respecto a la flota vehicular.

II. MARCO TEÓRICO

En la investigación de Briceño Yoel, con el título "Desarrollo de un sistema informático para mejorar la gestión de mantenimiento en la empresa Transportes Nacionales S.A."; la cual buscó explicar, cómo el desarrollo de un sistema informático pudo apoyar a la gestión de mantenimiento de la flota vehicular de la empresa Transportes Nacionales S.A, ya que es una herramienta que permitió dar soporte al proceso de la gestión de mantenimiento de la flota vehicular y también realizar mejores tomas de decisiones. Según la investigación, un sistema informático trae muchas ventajas, principalmente para la toma de decisiones en la gerencia, ya que permitió verificar el estado de mantenimiento de los vehículos, a través de informes y/o reportes suministrados de manera oportuna. En el alcance del trabajo se estimaron mejoras en medios técnicos, métodos de trabajo, materiales y resultados ya que se implementó un sistema que permitió el manejo de información y la evaluación del proceso. De igual manera se logró entender que la implementación de reportes y gráficos para la respectiva visualización y comprensión de los datos, permitió analizar la información de una manera óptima, lo que permitió ayudar a la toma de decisiones para la mejora de las actividades del proceso. Se concluyó que el desarrollo del sistema informático permitió mejorar el índice de conformidad general de la gestión de mantenimiento vehicular en un 17%, gracias a que tal sistema tuvo la capacidad de soportar las principales actividades del proceso de gestión de mantenimiento.

En la investigación de Castañeda Jackson & Gonzales Karim, con el título "Plan de mejora para reducir los costos en la gestión de mantenimiento de la empresa Transportes Chiclayo S.A."; en la cual se elaboró un plan de mejora con la finalidad de reducir los costos destinados a la gestión de mantenimiento de los vehículos de la flota vehicular de la empresa y así mismo fue planteada en la empresa. Para lo cual previamente se utilizaron técnicas para recolectar los datos, tales como observación, encuesta y entrevista, las cuales permitieron identificar a través de un Diagrama de Pareto los principales problemas, falta de políticas bien estructuradas, plan de mantenimiento, compromiso en los trabajadores, y paradas constantes

de la flota de vehículos. El plan de mejora de la gestión de mantenimiento permitió obtener un mejor control y registro de datos, mayor disponibilidad y confiabilidad de la flota de vehículos, para así disminuir los costos. Se concluyó que el tener adecuada gestión de mantenimiento en complemento con planes preventivos, programados y autónomos influye de manera significativa en los costos del mantenimiento de una flota, para beneficio de la empresa, ya que se disminuyó cerca al 50% de los costos de mantenimiento.

Asimismo, a nivel nacional en la investigación de Ricaldi Melissa, con el título "Propuesta para la mejora de la disponibilidad de los camiones de una empresa de transportes de carga pesada, mediante el diseño de un sistema de gestión de mantenimiento"; en la cual se buscó determinar cómo minimizar los tiempos del transporte de caña de azúcar para la empresa Interregional, a través del desarrollo de una propuesta de gestión de mantenimiento que permita optimizar la disponibilidad de los vehículos, y por ende permita realizar mayor cantidad de viajes, y entonces así se mejoren tanto los ingresos para la empresa de transportes como el nivel del servicio. Según la investigación, el principal problema son las demoras en los tiempos de transporte de caña de azúcar, lo que reduce el número de viajes y reduce los ingresos de dicha empresa. El 54% de las demoras se deben a los problemas mecánicos en los camiones, lo que genera paradas y en segundo lugar, el 29% de las demoras se debe a las intervenciones policiales. Debido al 54% de demoras por problemas mecánicos es que se analizó al área de mantenimiento, a través una auditoría, la cual tuvo como resultado un 51% de eficiencia por parte de esta área, lo cual se consideró como un bajo rendimiento. Se midió el desempeño del área con la referencia a los costos de mantenimiento con respecto a los ingresos, determinando el beneficio costo, y se obtuvo que un 6% de los ingresos es destinado a los mantenimientos. También se midió la disponibilidad, de la que se conoció que la flota cuenta con un 85% de disponibilidad y se mejoró el rendimiento de la flota vehicular en 5.45%, lo cual es aceptable. Las mejoras en el área de mantenimiento se realizaron en base a los datos de la flota. Se concluyó que la reducción de costos en

el mantenimiento de los camiones, puede tener como consecuencia el incremento de los costos de oportunidad que los camiones dejan de obtener por indisponibilidad por los problemas mecánicos.

En la investigación de Martín Macarena, con el título “Método para la gestión eficiente del combustible en flotas de vehículos con rutas fijas. Aplicación a una empresa de construcción”; en la cual se desarrolló una metodología para realizar una gestión de la flota de manera eficiente, enfocándose en el consumo de combustible, tal metodología fue aplicada basada en un estudio real de datos históricos. Según la investigación, la correcta implantación de un plan de mantenimiento de cada vehículo de la flota mejora su rendimiento, tanto en consumos como en días de disponibilidad del vehículo. La antigüedad de los vehículos también es un factor que implica en que el consumo aumente. Así mismo se determinó que existen veces en que es más beneficioso hacer un mantenimiento más exhaustivo para los vehículos mientras que otras veces la renovación de la flota sería una solución más rentable. También la conducción eficiente, disminuye el consumo de combustible y hace que la conducción sea más segura, de esta manera se disminuye el número de accidentes. Se concluyó que el consumo de combustible de vehículos se ve afectado por distintos factores, tales como la antigüedad, el mantenimiento, entre otros. Al gestionar eficientemente el consumo de combustible se puede reducir costos en los vehículos, y en relación a esto llevar a cabo un diagnóstico de la situación de la empresa en cada momento, para poder tomar las decisiones más adecuadas, logrando una reducción del porcentaje del costo por combustible en un 5%.

Asimismo, a nivel internacional en la investigación de Laureano Oscar, Garza Rodolfo & Mar Julio, con el título “Implementación de un sistema de gestión eficiente de flotas de transporte para la sostenibilidad económica en una empresa de transporte”; en la cual se buscó explicar cómo una metodología que consiste en la aplicación de una herramienta de gestión eficiente de flotas para medir los costos generados por el consumo de combustible y mantenimiento vehicular, puede contribuir a reducir los

costos operativos de la gestión de flota vehicular de la empresa Sociedad Cooperativa Trabajadores de Pascual. Según la investigación, los costos generados por la implementación del plan de mantenimiento, durante el periodo de Enero a Marzo del 2012, donde hubo una inversión de \$27,007.94 correspondientes al primer trimestre, sin embargo, durante el segundo trimestre solo se invirtió \$7,529.95, obteniendo una reducción de costos operativos del 72.12%, por lo tanto el plan de mantenimiento preventivo se dio con el fin de evitar inversiones con altos costos en mantenimientos correctivos. Se concluyó que se logró cumplir el objetivo planteado ya que el porcentaje de costos en mantenimiento, lavado, infracciones y neumáticos, se redujeron en 6.15%, 2%, 8% y 3.31% respectivamente, en la relación a los costos de operación en la administración de la flota de transporte.

En la investigación de PRNewswire, denomina a SageQuest como un software líder de gestión de flotas basado en GPS y software de business intelligence (SaaS). El ahorro de combustible, menos accidentes y el servicio oportuno al cliente son solo una fracción de los beneficios que se obtienen por el uso de dicho software. SageQuest permite a las organizaciones que cuentan con una flota vehicular, enfocarse en otras áreas de misión crítica de su negocio. Así mismo se investigó que Fleetmatics Group PLC es un proveedor global líder de soluciones de gestión de flotas para pequeñas y medianas empresas, entregado como software como servicio (SaaS). Dichas soluciones permiten a las empresas enfrentar los desafíos asociados con la administración de flotas locales y mejorar la productividad de sus fuerzas de trabajo móviles, extrayendo mediante business intelligence datos para procesar el comportamiento de conductores y vehículos en tiempo real e históricos.

Las soluciones de flotas ofrecen visibilidad sobre la ubicación del vehículo, el uso de combustible, la velocidad y el kilometraje, y otros conocimientos sobre su fuerza de trabajo móvil, lo que les permite reducir los costos operativos y de capital, así como aumentar los ingresos. Fleetmatics sirve a más de 17,000 clientes, rastreando más de 330,000 vehículos en todo el

mundo. Las soluciones de Fleetmatics Group se comercializan bajo las marcas Fleetmatics y SageQuest..

Según la investigación de Ryan E.; en la gestión de flotas, el vehículo no puede separarse de su carga ni de su conductor. Además, se está convirtiendo cada vez más en un componente que también se ha vuelto inseparable de la tecnología, específicamente en la business intelligence o como componente analítico. En el caso de las empresas de transporte, la gestión de flotas es claramente su columna vertebral. Sin embargo, esto también es válido para los vehículos de reparto de muchas empresas, los vehículos de personal o los vehículos de construcción. La administración de flotas comenzó como una industria en Sudáfrica a fines de la década de 1970, siendo poco más que una tarjeta de combustible / mantenimiento y un sistema de informes mensuales. Evolucionó junto con la tecnología, pero recientemente, como es el caso en muchas otras industrias, ha habido una ola de nueva legislación y requisitos de gobierno que requieren la automatización de la gestión de la flota simplemente para hacer frente. Se ha encontrado que incluso en una cadena de suministro eficiente, todavía hay hasta un 10% de costo que se puede reducir. Por lo tanto, la cadena de suministro eficiente trata de administrar la red y toda la cadena de valor.

También se ha encontrado IMS Fleet Intelligence el cual es otro ejemplo que permite a las organizaciones controlar mejor el fraude de combustible, los gastos por horas extra y el uso no autorizado de vehículos, al tiempo que reduce el riesgo de incidentes de daños a través del entrenamiento de conductores y la gestión del comportamiento. También permite a las aseguradoras de líneas comerciales obtener un nuevo flujo de ingresos y valor orientados a la flota, estos son servicios agregados que construyen la lealtad del cliente.

GE Capital Fleet Services es una de las compañías de gestión de flotas más grandes del mundo, con más de 1 millón de automóviles y camiones comerciales bajo gestión de arrendamiento y servicio. Tiene operaciones en los Estados Unidos, Canadá, México, Europa, Japón, Australia, Nueva

Zelanda y Brasil. GE Capital, con activos por más de US \$ 345 mil millones, es una compañía global de servicios financieros diversificados con 28 negocios especializados. Una subsidiaria de propiedad total de General Electric Company, GE Capital, con sede en Stamford, Connecticut, ofrece administración de equipos, financiamiento de mercado intermedio y especializado, seguros especiales y una variedad de servicios al consumidor, como alquiler de automóviles, hipotecas y tarjetas de crédito, a empresas y personas de todo el mundo. GE es una empresa de fabricación, tecnología y servicios con operaciones en todo el mundo. GE Capital Fleet Services es una de las compañías de gestión de flotas corporativas líderes en el mundo, con más de \$ 9 mil millones en activos y más de 1 millón de automóviles y camiones bajo la gestión de alquileres y servicios. Los servicios de la compañía se centran en todo el ciclo de vida de un vehículo: financiamiento; selección y adquisición; Impuestos, títulos y licencias, programas de gastos operativos tales como servicios de gestión de mantenimiento y prevención de accidentes, tarjeta electrónica de combustible, renovación de registro, informes de gastos y gestión, compras de descuentos y subcontratación de gestión de flotas.

Según PR Newswire; PHH Arval, una subsidiaria de PHH Corporation (NYSE: PHH), informó los resultados positivos de los clientes que realizaron pruebas beta a PHH Onboard (SM), su servicio de flota telemática. El servicio se desarrolló combinando la experiencia en consultoría de PHH, datos de la flota en el Data Warehouse de PHH y la tecnología de gestión inalámbrica de la flota de Networkcar. De los doce clientes y cientos de vehículos de la flota que participan en la prueba beta, los primeros resultados indicaron que la telemática de la flota redujo drásticamente los perfiles de riesgo de la flota, redujo los costos de mantenimiento y combustible, y mejoró la productividad.

Por otro lado, según Tobias S., la empresa Norfolk Southern, es la segunda flota vehicular de mayor propiedad en América del Norte. Su objetivo es poner la flota de automóviles en línea con las necesidades del cliente y al mismo tiempo brindar los beneficios de la mejora. La parte más visible de

ese esfuerzo ha sido la decisión de disponer de 12,000 autos excedentes. A mediados de año, NS tenía compromisos por valor de 33 millones de dólares para 10,000 autos. También se informó sobre los avances en los sistemas de gestión de flotas que han llevado a la generación de más información, de manera más oportuna. La gestión de automóviles está aprovechando eso mediante el desarrollo de aplicaciones de almacenamiento de datos que transmitieron información de la flota a cada departamento en NS que tenga un impacto en la utilización del automóvil. Con las herramientas de almacenamiento de datos, se pudieron capturar más rápidamente los tiempos de ciclo y medir el rendimiento en comparación con el plan de los automóviles. El almacén de datos permitió visualizar de manera instantánea el estado del inventario de vagones de carga y analizar la información de alquiler de automóviles de manera rápida y efectiva. También se desarrollaron otras herramientas para facilitar el proceso de pedido de automóviles y para desarrollar un modelo de planificación de flotas. La utilización de la flota ha mejorado de 53.2% a 69.2%.

Así mismo Asset Intelligence ha basado su nuevo servicio de análisis empresarial en una plataforma Teradata Data Warehouse Appliance, Teradata Corporation. La cual permitió ayudar a los transportistas de carga por carretera a eliminar los puntos ciegos en sus operaciones y a impulsar una mejor productividad y rendimiento de la flota. Asset Intelligence, con la adición de un almacén de datos impulsado por la plataforma Teradata Data Warehouse Appliance, espera ofrecer un análisis de datos más rápido e inteligente, permitiendo a sus clientes reaccionar más rápidamente a los desarrollos en tiempo real del estado y condición de los bienes. Además, ofrece soluciones telemétricas para el seguimiento y la gestión de flotas de remolques y contenedores a través de su sistema VeriWise. Teradata Corporation es un proveedor de software de bases de datos, almacenamiento de datos empresariales, dispositivos de almacenamiento de datos, consultoría y analítica empresarial.

Así mismo la herramienta de inteligencia empresarial (BI) para el sistema de gestión de flotas TruckMate; permitió dar una ventaja de en la industria

de camiones; generando un impacto de BI positivo en la rentabilidad de los operadores de camiones.

Sistema de Soporte de Decisiones

Es un tipo de sistema similar a un sistema de información administrativa tradicional, esto es debido a que existe una obvia dependencia de una fuente de datos, por ejemplo, una base de datos. El sistema de soporte de decisiones está orientado a poder brindar un soporte para la toma de decisiones, pero la decisión siempre es tomada por algún usuario o responsable de negocio. Los sistemas de soporte de decisiones están enfocados a transmitir conocimiento del negocio a la persona o grupo de personas que requieran de una correcta gestión de la información para cumplir con sus necesidades y objetivos, permitiendo obtener ventajas competitivas. Este tipo de sistema es considerado como la mejor opción para ser utilizado con la business intelligence (Kendall & Kendall, 2014).

Business Intelligence

La Inteligencia de negocios o mejor conocido Business Intelligence (BI), es un amplio conjunto de soluciones de TI que incluye herramientas para recolectar, analizar y reportar la información de una organización, permitiendo brindar soporte a la toma de decisiones, a través de soluciones creativas para una variedad de problemas. Las organizaciones invierten en herramientas de BI para ganar ventajas competitivas, teniendo el objetivo de tomar decisiones efectivas mientras se reducen los riesgos. El BI tiene la capacidad de ayudar a mejorar el rendimiento de casi toda la diversidad de industrias y sus aplicaciones (Maheshwari, 2015).

Data Mart

Son almacenes de datos con un enfoque de implementación por departamentos en una organización, sin preocupación de compartir e integrar la información de toda la organización. Por lo general, un departamento individual identifica sus requerimientos de datos y se construye una base de datos que satisfaga sus necesidades

departamentales. Esta información se encuentra aislada en un Data Mart. Por otro lado, es común que varios departamentos estén interesados en la misma información de los eventos del proceso central de una organización, pero debido a que un departamento no tiene acceso a Data Mart inicialmente construido por otro departamento, se procede por construir un Data Mart similar para sus necesidades de información, generando una solución que contenga datos similares, pero ligeramente diferentes. El problema ocurre cuando los usuarios de estos departamentos debaten sobre el desempeño organizacional basados en informes obtenidos de sus respectivos repositorios, y ninguno de los números coinciden debido a las diferentes reglas de negocio que maneja cada uno. Entonces un Data Mart puede ser válido cuando no se requiere coordinación de datos en la organización, además tiene un desarrollo a corto plazo. Pero las extracciones iterativas de las mismas fuentes de datos operacionales y almacenamiento de datos de manera redundante, son ineficaces y perjudiciales a largo plazo (Kimball & Ross, 2013).

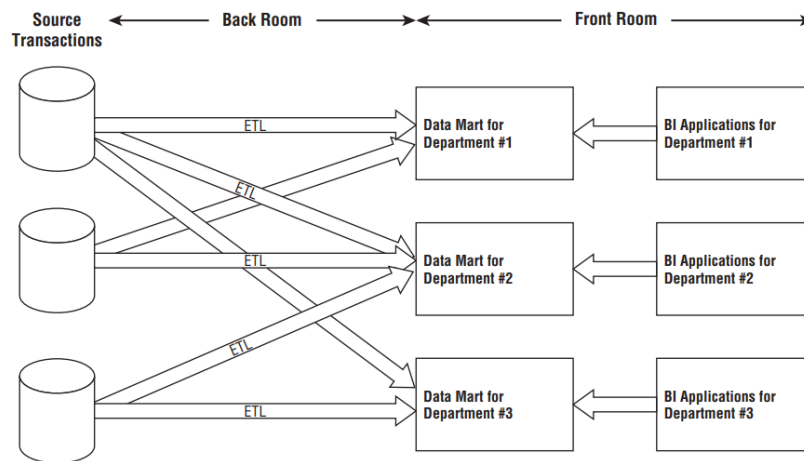


Figura N°1 Arquitectura Data Mart

Fuente: Kimball R. & Ross M. The Data Warehouse Toolkit.

2.1.1.1 Data Warehouse

El Data Warehouse (DW) es una colección de bases de datos integradas de toda una organización, diseñado para respaldar el

soporte de decisiones. Los datos son extraídos de una Base de Datos operacional para responder a un conjunto de consultas particulares. Estos datos combinados con otros, se acumulan hasta obtener una granularidad consistente y son cargados en un almacén de datos llamado DW, el cual está separado física y funcionalmente de la base de datos. Entonces un DW es una simple versión de una base de datos operacional, con el propósito de abordar las necesidades de reportes y toma de decisiones. La creación de un DW para una organización representa una significativa inversión en tiempo y esfuerzo, por lo que, debe ser constantemente actualizada para que sea de utilidad. Un DW es compatible con los reportes del negocio y las actividades de minería de datos (Maheshwari, 2015).

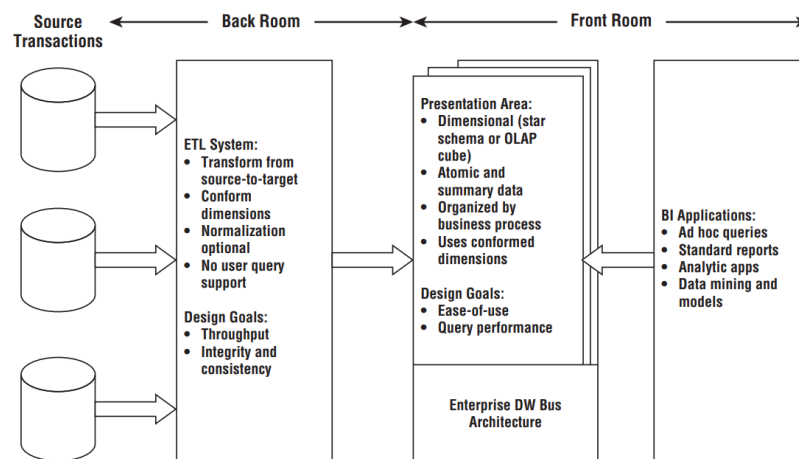


Figura N°2 Arquitectura Data Warehouse

Fuente: Kimball R. & Ross M. The Data Warehouse Toolkit.

Extracción, Transformación y Carga (ETL)

Es el proceso que sirve para poblar un Data Warehouse con datos de una buena calidad. En la extracción, los datos se deben extraer de muchas fuentes de datos de forma regular. Para la transformación, los datos extraídos deben estar alineados de manera uniforme por campos clave, por lo que se debe limpiar cualquier irregularidad o valores perdidos en los datos. Los datos necesitan estar al mismo nivel de granularidad. Se calculan los campos tales como, por ejemplo, los totales de ventas diarias, y toda la información debe ser

transformada al mismo formato que la tabla central del Data Warehouse. Por último, en la carga, los datos transformados se deben cargar en el Data Warehouse. Este proceso ETL debe ejecutarse a una frecuencia regular en un ciclo. En las transacciones diarias los datos pueden extraerse de los ERP o de cualquier otra herramienta que centralice la información, transformarse y cargarse en la base de datos; para que el Data Warehouse siempre se encuentre actualizado. El trabajo de ETL está usualmente automatizado utilizando scripts de programación para la actualización periódica del Data Warehouse (Maheshwari, 2014).

Modelado Dimensional

El modelado dimensional es una técnica que permite simplificar las bases de datos. La simplicidad es un factor clave porque garantiza que los usuarios puedan comprender con facilidad los datos, y permite que una aplicación o software acceda y entregue los resultados de una manera rápida y eficiente. Los modelos dimensionales son establecidos en los sistemas de gestión de base de datos relacionales, pero no poseen una estructura normalizada. El modelo dimensional contiene la misma información que un modelo que se encuentra normalizado, pero almacena los datos en un formato que permite mejorar la comprensión del usuario, el rendimiento de las consultas y la capacidad de cambios (Kimball & Ross, 2013).

Star Schema y OLAP

Los modelos dimensionales implementados en entornos de bases de datos relacionales se denominan esquemas estrella. Por otro lado, los modelos dimensionales que se encuentran implementados en entornos de bases de datos multidimensionales se denominan cubos de procesamiento analítico en línea (OLAP). Si un entorno de Business Intelligence y/o Data Warehouse posee esquemas estrella o cubos OLAP, está aprovechando los conceptos dimensionales.

Tanto los esquemas estrella como los cubos OLAP poseen un diseño lógico común con dimensiones entendibles; sin embargo, la implementación física es lo que cambia. Cuando los datos son cargados en un cubo OLAP, se almacenan e indexan utilizando formatos y técnicas que están diseñados para datos dimensionales. Entonces el rendimiento y las tablas de resumen a menudo son creados y administrados por el motor de cubos OLAP. En consecuencia, los cubos ofrecen un rendimiento en la consulta superior debido a las tablas de resumen, la indexación y otras optimizaciones. Los usuarios que requieran consultar la información pueden ir añadiendo o eliminando atributos de sus análisis con un rendimiento excelente sin realizar nuevas consultas. Los cubos OLAP proveen funcionalidades analíticamente robustas que superan a las funcionalidades de SQL. La desventaja es que existe un alto consumo del rendimiento de los procesadores por las funcionalidades y capacidades de los cubos OLAP, sobre todo con grandes conjuntos de datos. Aunque la tecnología OLAP mejoran continuamente, son los esquemas estrellas en donde preferentemente se carga la información detallada; y los cubos OLAP se complementaran a partir del esquema estrella (Kimball & Ross, 2013).

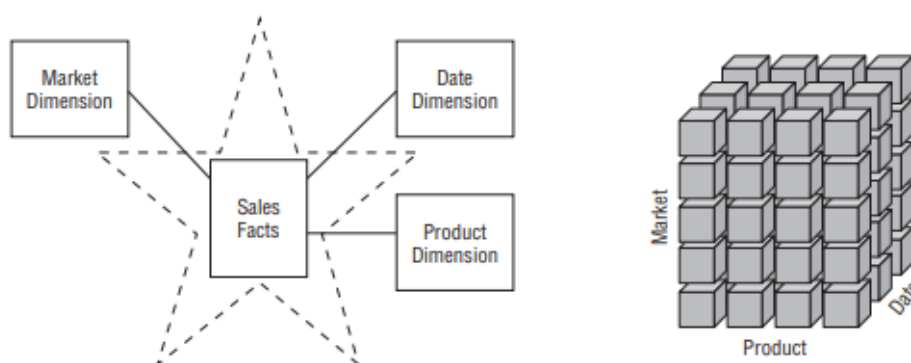


Figura N°3 Star Schema & OLAP

Fuente: Kimball R. & Ross M. The Data Warehouse Toolkit.

Visualización de datos

Una visualización de datos transmite propiedades únicas del conjunto de datos que representa. Las visualizaciones ayudan a dar una vista rápida

del conjunto de datos a los usuarios, y pueden ser implementados en alguna aplicación o software para manejar la información almacenada (Fry, 2008).

Son representados en gráficos que proporcionan un enfoque excelente para la exploración de los datos y son esenciales para presentar los resultados. Por otro lado, aunque los gráficos se han utilizado ampliamente en las estadísticas durante mucho tiempo, no existe una teoría sustancial que sustente tal el tema (Chen, Härdle & Unwin, 2008).

El objetivo principal de la visualización de datos es hacer que el lector abstraiga rápidamente los datos, junto con posibles tendencias, relaciones, entre otras. Por esta razón, se deben tomar las visualizaciones con importancia y asegurar que son lo más efectivas posibles. Se tienen los siguientes tipos básicos de gráficos: gráficos de dispersión, gráficos de líneas, gráficos de barras, histogramas y gráficos circulares (Ozdemir, 2016).

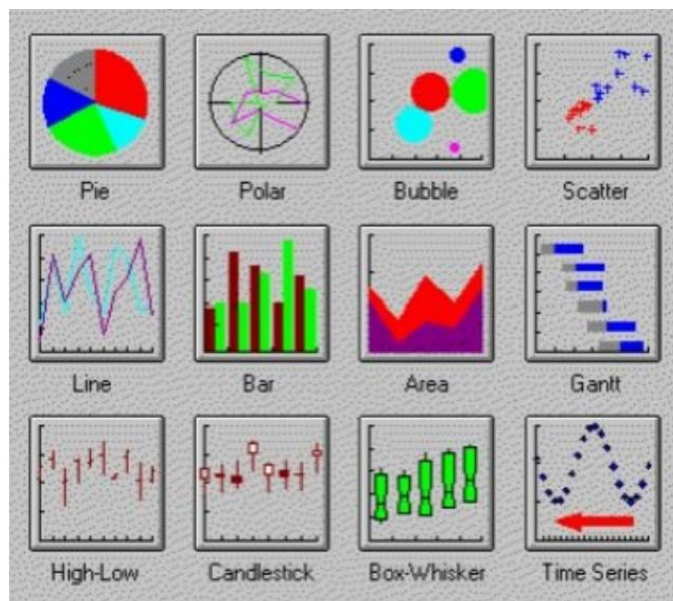


Figura N°4 Tipo de gráficos para visualización de datos

Fuente: Maheshwari A. Data Analytics Made Accesible.

Data Mining

La Minería de Datos o Data Mining (DM) es el arte y ciencia de descubrir conocimiento, ideas y patrones útiles en los datos. Hay muchas técnicas,

simples o complejas, que ayudan a encontrar patrones. Un patrón es un diseño o modelo que ayuda a comprender algo, también posee cualidades como: describe una situación de forma precisa, es ampliamente aplicable y puede ser descrito de una manera simple. Los datos sobre el pasado pueden revelar patrones que pueden proyectarse en el futuro, ya que los datos pasados pueden tener valor predictivo en situaciones complejas, especialmente en donde el patrón no es fácilmente visible sin una técnica de modelado. El DM es un campo multidisciplinario que utiliza distintas técnicas de una variedad de campos. Utiliza técnicas de modelado y análisis provenientes de las áreas de estadística e inteligencia artificial. También aprovecha el conocimiento de la toma de decisiones en el campo de la gestión empresarial (Maheshwari, 2015).

Big Data

Big data es un término amplio para una colección de grandes conjuntos de datos; siendo tan complejos que resultan difícil procesarlos con herramientas tradicionales de administración de datos. Big data es más valioso cuando se analizan los datos en conjunto. Sin embargo, se necesitan habilidades y herramientas especializadas en administrar dichos conjuntos de datos extremadamente grandes. Los datos masivos pueden ser usados para descubrir nuevas perspectivas del negocio, desde un enfoque de 360 grados sobre una situación que puede permitir una nueva perspectiva, nuevos modelos de negocio y nuevas soluciones potenciales. Los datos se han convertido en el nuevo recurso natural. Las organizaciones con grandes cantidades de datos, pueden escoger entre ser enterradas bajo la avalancha de datos, o pueden optar por usarlos para obtener una ventaja competitiva. Big Data puede ayudar a determinar la calidad de investigaciones, prevenir enfermedades, enlazar hechos legales, combatir la delincuencia, determinar las condiciones del tráfico en la carretera en tiempo real y detectar tendencias u oportunidades comerciales (Maheshwari, 2015).

En el año 2000, había 800 000 Petabytes de datos en el mundo. Se espera que para el año 2020, existan a 35 Zettabytes. Big Data es grande, rápido, desestructurado, y de muchos tipos (Maheshwari, 2015).

Data Science

La ciencia de los datos implica principios, procesos y técnicas para comprender los fenómenos de los datos, a través de un análisis automatizado. Pretende analizar cómo mejorar la toma de decisiones, basándose en los datos, ya que esto generalmente es de interés directo para las organizaciones; en el contexto de como los procesos están relacionados con los datos. La ciencia de datos abarca aspectos de procesamiento de datos que están ganando cada vez más atención en los negocios. Esto pone en claro de que cada vez más, las decisiones comerciales las toman automáticamente los sistemas informáticos. Diferentes industrias han adoptado la toma de decisiones automática a diferentes ritmos. Las industrias de finanzas y telecomunicaciones fueron las primeras en adoptarlo, lo que les permitió agregar y hacer modelos de datos a gran escala (Big Data), así como la aplicación de los modelos de aprendizaje (Machine Learning), obteniendo resultados para la toma de decisiones (Provost & Fawcett, 2013).

Actualmente la profesión en el campo de la ciencia de los datos (Científico de Datos) tiene una alta demanda y la oferta no está a la altura, y las organizaciones requieren de los datos y el análisis de estos para intentar predecir el futuro en las distintas áreas y así ganar ventajas competitivas, además de generar ahorros y beneficios directos con los consumidores o proveedores (Davenport & Patil, 2012).

Machine Learning

El término aprendizaje automático o machine learning se refiere a la detección automática de patrones significativos en los datos. En las últimas dos décadas se ha convertido en una herramienta común en casi cualquier tarea que requiera la extracción de información de grandes conjuntos de

datos. Estamos rodeados de una tecnología basada en el aprendizaje automático: los motores de búsqueda aprenden a ofrecernos los mejores resultados (al colocar anuncios rentables), el software antispam aprende a filtrar nuestros mensajes de correo electrónico y las transacciones de tarjetas de crédito se aseguran mediante un software que aprende cómo para detectar fraudes. Las cámaras digitales aprenden a detectar caras y las aplicaciones de asistencia personal en los teléfonos inteligentes tiene la capacidad de reconocer los comandos de voz. Los automóviles están equipados con sistemas de prevención de accidentes que se crean utilizando algoritmos de aprendizaje automático. El aprendizaje automático también se usa ampliamente en aplicaciones científicas como la bioinformática, la medicina y la astronomía. Una característica común de todas estas aplicaciones es que, en contraste con los usos más tradicionales de las computadoras, en estos casos, debido a los patrones con alta complejidad que necesitan ser detectados, un programador humano no puede proporcionar una especificación de cómo se deben ejecutar tales tareas. Muchas de nuestras habilidades se adquieren o refinan aprendiendo de nuestra experiencia. Las herramientas de aprendizaje automático se ocupan de dotar a los programas con la capacidad de aprender y adaptarse (Shai & Shai, 2014).

Modelos de Aprendizaje

Dado que el aprendizaje implica una interacción entre el alumno y el entorno, uno puede dividir las tareas de aprendizaje de acuerdo con la naturaleza de esa interacción. (Shai & Shai, 2014).

Supervisado

Se enseña o entrena al algoritmo con datos que ya están estructurados y que han sido establecidos con etiquetas. Mientras más grande sea el conjunto de datos el algoritmo aprende más. Una vez entrenado, se le brindan nuevos datos y sin etiquetas, entonces el algoritmo de aprendizaje utiliza la

experiencia pasada que adquirió durante la etapa de entrenamiento para poder predecir un resultado. (Shai & Shai, 2014).

No Supervisado

El algoritmo es entrenado con el uso de un conjunto de datos sin ninguna etiqueta. Esto es con el fin de que el algoritmo pueda encontrar por si solo patrones que le ayuden a comprender el conjunto de datos. (Shai & Shai, 2014).

Aprendizaje por refuerzo

El algoritmo aprende comprendiendo su entorno. Los datos de entrada son la retroalimentación que obtiene de su entorno en consecuencia a sus acciones. Por lo tanto, el algoritmo aprende en a base de prueba y error. (Shai & Shai, 2014).

Algoritmos de Machine Learning

Regresión Lineal

Se utiliza para estimar valores reales basados en variables. Se establece la relación entre las variables independientes y dependientes con el fin de ajustar una línea recta con respecto a los puntos, referentes a las variables. Esta línea ajustada es la línea de regresión y está representada por una ecuación lineal.

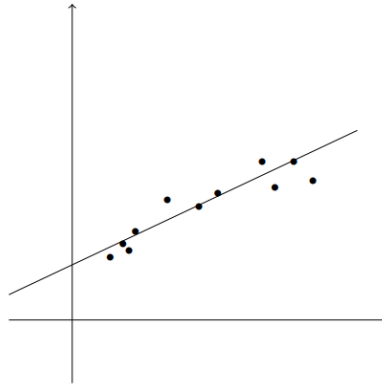


Figura N°5 Regresión Lineal

Fuente: Shai S. & Shai.B. Understanding Machine Learning: From Theory to Algorithms

Regresión Logística

Es utilizado para clasificar; se ajusta el modelo lineal a la probabilidad de que una cierta situación ocurra y luego se utiliza una función para crear un umbral en el cual se especifica el resultado de la situación. Este modelo utiliza la función logística.

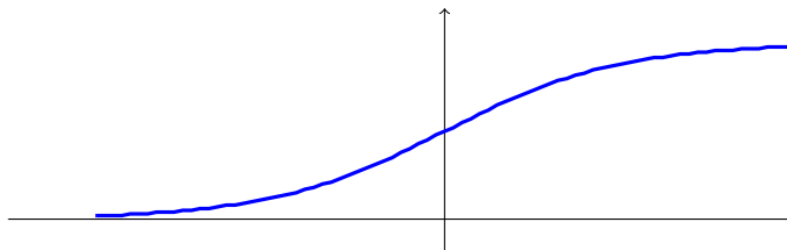


Figura N°6 Regresión Logística

Fuente: Shai S. & Shai.B. Understanding Machine Learning: From Theory to Algorithms

Árboles de Decisión

Son diagramas con construidos de manera lógica, muy similares a los algoritmos de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren sucesivamente, para solucionar un problema. Los árboles de decisión consisten en nodos

internos, nodos terminales y ramas que interrelacionan los nodos. Los nodos interiores contienen una prueba de un atributo, y cada rama representa un valor diferente al del atributo. Finalmente, en el nodo terminal se crea una segmentación de los datos.

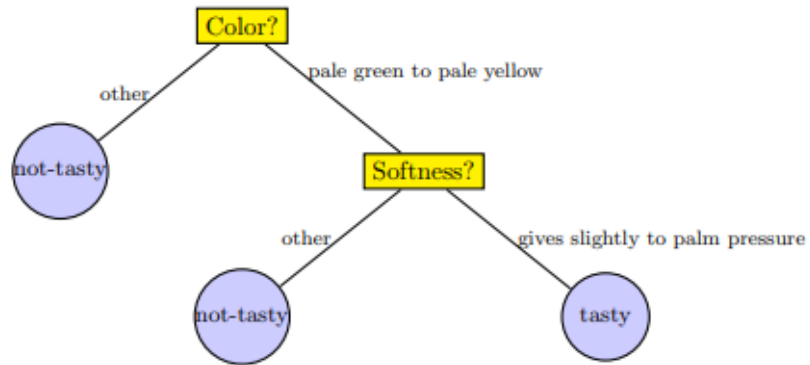


Figura N°7 Decision Tree

Fuente: Shai S. & Shai.B. Understanding Machine Learning: From Theory to Algorithms

Random Forest

Un concepto simple es definirlo como un bosque de árboles de decisión. Se construyen árboles de decisión muy poco profundos, y luego toma la clase de cada árbol. Esta idea es muy útil en Machine Learning.

Support Vector Machine (SVM)

Se centra en la clasificación basándose en encontrar un plano que separe los grupos dentro de los datos de la manera más efectiva. La separación permite que la elección del plano amplíe el margen existente entre los puntos más cercanos en el plano, estos puntos se denominan vectores de soporte.

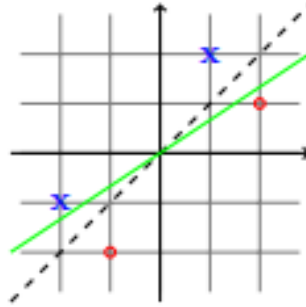


Figura N°8 Support Vector Machine

Fuente: Shai S. & Shai.B. Understanding Machine Learning: From Theory to Algorithms

K Near Neighbors (KNN)

Es un método de clasificación, que estima el valor de la probabilidad de que un elemento X pertenezca a una clase en particular a partir de la información proporcionada por un conjunto de prototipos. La regresión KNN se calcula tomando el promedio del punto K más cercano al punto X.

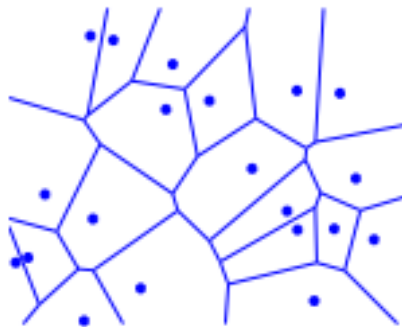


Figura N°9 K Near Neighbors

Fuente: Shai S. & Shai.B. Understanding Machine Learning: From Theory to Algorithms

K-Means

Es uno de los algoritmos de clasificación más conocidos así como una de las técnicas de aprendizaje no supervisado más conocidas. K-means es en realidad un algoritmo

relativamente simple que funciona para reducir al mínimo la suma de las distancias cuadradas desde la media dentro del grupo. Para hacer esto establece primero un número previamente especificado de conglomerados, K, y luego va asignando cada observación a al grupo más cercano de acuerdo a su media.

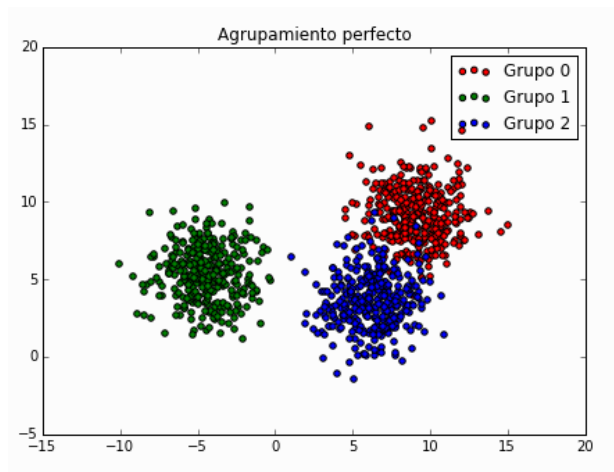


Figura N°10 K-Means

Fuente: Shai S. & Shai.B. Understanding Machine Learning: From Theory to Algorithms

Contexto tecnológico

Lenguaje de Programación

Java

Java es un lenguaje de programación orientado a objetos y una plataforma que fue comercializada por primera vez en 1995 por Sun Microsystems, y fue desarrollado originalmente por James Gosling. Java es ampliamente aceptado al ser gratuito, rápido, seguro y fiable. En la actualidad muchas aplicaciones de escritorio, web o móvil están desarrolladas en Java. Desde dispositivos portátiles hasta centros de datos, desde consolas de videojuegos hasta súper computadoras, desde teléfonos móviles hasta en Internet, Java está en todas partes. Siempre se realizan continuas actualizaciones que

contienen importantes mejoras para el rendimiento, estabilidad y seguridad de las aplicaciones Java. El Java Runtime Environment (JRE) es todo lo que se necesita para ejecutar Java en cualquier plataforma; está formado por el Java Virtual Machine (JVM), clases del núcleo de la plataforma Java y bibliotecas de la plataforma Java de soporte (Oracle, 2014).

El Java Development Kit (JDK) es el conjunto de herramientas necesarias para el desarrollo de aplicaciones en Java (Oracle, 2018).

Python

Python es un lenguaje de programación interpretado y posee sintácticamente una baja curva de aprendizaje. La sintaxis de Python es simple y fácil de comprender, además de ser un lenguaje de tipado dinámico. Cuenta con estructuras de datos eficientes, y está enfocado a minimizar las dificultades de la sintaxis, pero continuar siendo efectiva como lo es en la programación orientada a objetos. Python permite un desarrollo rápido y versátil de aplicaciones en diversas áreas y sobre la mayoría de las plataformas. El intérprete de Python y su extensa librería están a libre disposición en forma binaria y de código fuente, y puede distribuirse libremente (Python Documentation, 2017).

R

R es un lenguaje de programación de alto nivel e interpretado, de distribución libre, bajo Licencia GNU, y se mantiene en un entorno para el cómputo gráfico y estadístico. El término entorno pretende caracterizarlo como un sistema totalmente y coherente, y no solo una acumulación gradual de herramientas que sean poco flexibles, como suele ser con otros softwares de análisis de datos. El hecho que R sea un

lenguaje y un sistema, es porque forma parte de la filosofía de creación, por lo que, R es un sistema estadístico y un entorno en el que se aplican técnicas estadísticas. Este software se puede ejecutar en distintas plataformas como Linux, Windows, MacOS, e incluso en consolas PlayStation 3 (Santana & Farfán, 2014).

WEKA

La herramienta WEKA (Waikato Environment for Knowledge Analysis) provee una colección de implementaciones de algoritmos de aprendizaje automático y preprocesamiento de datos. Fue diseñado para probar rápidamente distintos métodos en distintos conjuntos de datos de manera flexible. Brinda un gran soporte para el proceso de extracción de datos; la preparación de datos, evaluación de los esquemas de aprendizaje a nivel estadístico, visualización y el resultado del aprendizaje. La herramienta brinda una interfaz amigable para que los usuarios puedan fácilmente comparar diferentes métodos e identificar aquellos que son más apropiados para su problemática. WEKA fue desarrollado en la Universidad de Waikato en Nueva Zelanda. El software está desarrollado en Java y distribuido bajo los términos de la GNU General Licencia pública. Se ejecuta en distintas plataformas y ha sido probado en Linux, Windows y MacOS (Witten, Frank, Hall & Pal, 2017).

Gestión de Flota

La gestión de flota es la administración de un conjunto de vehículos y de sus procesos relacionados, que tiene como finalidad para la organización, el prestar un servicio a un cliente o en el apoyo de las actividades en la misma organización, de la manera más eficiente, para el cumplimiento del nivel de servicio y el costo, establecido en la organización. La gestión de la flota de vehículos es un factor de vital importancia para apoyar al

desarrollo y correcto cumplimiento de las estrategias y objetivos de una organización, por lo que debe ser implementada en base a las directrices de la organización. Hay organizaciones en las que la gestión de su flota de vehículos es su principal proceso de negocio y lo que les genera valor, y por lo tanto están orientadas a la prestación de servicios, tales como traslado de turistas, traslado de mercadería, alquiler de vehículos, servicios de taxis servicios de mensajería, servicios de camiones de basura, entre otros. En cambio, para otras organizaciones es una actividad que brinda soporte a su núcleo del negocio y su correcto desempeño, o tal vez para el traslado del personal directivo, asignándoles un vehículo para su uso particular, tales como empresas comerciales, aeropuertos, supermercados u organizaciones del gobierno. Cada flota tiene características, variables y factores distintos, por lo que cada flota de vehículos es única y diferente, y entonces se debe determinar la mejor solución que sea más factible para la gestión de la misma. (Fernández, 2016).

Actividades de la Gestión de Flota

Política de la gestión de flota de vehículos

Sirve para que se establezcan directrices que sirvan para la planificación que determine el buen uso y manejo de los recursos, con el objetivo de que se desarrolle la estrategia de la organización y se cumplan satisfactoriamente, dando mejoras en el nivel de servicio y costo. La dirección de una organización tiene que diseñar, desarrollar e implementar la política de la gestión de la flota a partir de la estrategia de la organización, y es su deber el velar por el cumplimiento de las directrices, las cuales deben de ser transmitidas de manera oportuna a todas las personas involucradas. La ausencia de una política bien establecida, muchas veces es el origen de una mala gestión de la flota de vehículos (Lockhart, 2016).

Cálculo del número óptimo de vehículos de la flota

El cálculo del número de vehículos de una flota es complicado, debido a todos los factores que se involucran, como lo son el establecer el nivel de servicio y el costo respectivo, la planificación de las mejores rutas si previamente se determinaron cuáles eran, la clase de vehículos a ser utilizados y las capacidades que estos poseen para que puedan ser utilizada al 100% o determinar la mejor localización para establecer una base de operaciones (Fernández, 2016).

Adquisición de los vehículos de la flota

La adquisición de los vehículos es establecida previamente por una política que define cual es la normativa de la adquisición, entre estas se deben determinar cuál es la modalidad para la adquisición de vehículos, los tipos de vehículo, las marcas, los modelos, el equipamiento, el tipo de combustible, el tipo de seguro y los proveedores o en este caso las concesionarias de los vehículos. Hay varias opciones de como adquirir los vehículos, tales como, adquiriéndolos como propiedad, mediante renting o leasing, a través de la financiación de crédito o haciendo uso de carsharing, la elección básicamente depende de la estrategia planteada por la organización (Fernández, 2016).

Operación de los vehículos de la flota

Es lo relacionado con gestión de los recambios, gestión de las averías, gestión del mantenimiento, gestión de los neumáticos, reducción y control de consumo del combustible, la reducción y prevención de los accidentes de tráfico o utilización de talleres que sean externos o propios. La organización se encarga de establecer directrices referentes a la utilización de los vehículos, consumo de combustible, mantenimiento, manera de como conducir, la seguridad y cuál es la medida a tomar cuando ocurre un accidente o avería, se

deben poner en conocimiento a cada conductor y/o responsable de un vehículo, sobre estas directrices para ser acatadas. El principal objetivo de la correcta operación de los vehículos es que los vehículos se encuentren con el mayor tiempo disponible para cumplir el nivel de servicio establecido, y con los menores costos variables de operación que sean posibles (Fernández, 2016).

Retirada de los vehículos de la flota del servicio

Tiene la finalidad de retirar los vehículos de flota en base a una serie de factores determinados por la organización, para proceder con la respectiva renovación de la flota o bloque de vehículos, se debe escoger a los proveedores de venta, el canal de venta o establecer el tiempo máximo para vender los vehículos. El tiempo máximo de la utilización de un vehículo es definido por su periodo de renovación, y una vez cumplido el periodo se procede a ser renovado o retirado del servicio y vendido. El objetivo de la retirada y renovación de vehículos es lograr vender los vehículos en su debido momento y al máximo precio posible, y así utilizar este ingreso para la adquisición de nuevos vehículos u otras necesidades de la flota requeridas. Se debe tener en cuenta que desde que el vehículo es retirado del servicio hasta que sea vendido se generan costos de transporte y almacenaje que deberían ser minimizados. El vehículo debe ser vendido en óptimas condiciones, y deberá ser reparado de ser requerido, puede ser necesario certificar el estado del vehículo por una empresa externa. Si el vehículo se adquiere bajo la modalidad de renting o leasing, este debe ser devuelto durante o a la finalización del contrato. La organización es la que establece la política de renovación del vehículo, la cual puede ser, por ejemplo; cuando se alcance un determinado kilometraje, la antigüedad del vehículo o lo que antes se cumpla antes, entre otros (Fernández, 2016).

Actividades de la Operación del Vehículo

Gestión del Mantenimiento

Tiene mayor importancia, porque si no se realiza un adecuado mantenimiento en un lapso de tiempo y también el hecho de no seguir las indicaciones del fabricante, tendrían estas consecuencias: problemas con la correa de distribución o el motor por no haber renovado el aceite; no detectar y prever elementos defectuosos y gastados o posibles averías afectando la integridad del vehículo generando problemas de seguridad, y por lo tanto se generen un mayor número de accidentes, averías, emisiones contaminantes y consumo de combustibles. Así mismo también se incrementan las posibilidades de incrementar los costos de la organización, siendo una desventaja para esta. Se debe establecer un sistema para planificar satisfactoriamente el mantenimiento vehicular, por lo que es necesario tener conocimiento del kilometraje o la actividad en horas del vehículo regularmente (Gastalver, 2017).

Gestión de los Repuestos

En esta actividad se deben establecer los tipos de repuestos que serán utilizado por la flota vehicular, suministrados por proveedores o por el mismo fabricante. Aproximadamente el 80% de componentes que pertenecen a un vehículo son suministrados por proveedores externos al fabricante, por lo que no existe diferencia sustancial en estos repuestos suministrados. La principal ventaja de adquirir un repuesto de un proveedor externo al fabricante es que es más económico, y se puede obtener hasta un ahorro de costos del 40%, pero se debe analizar la existencia o especificidad del repuesto a ser suministrado basándose en la especialización del repuesto (Fernández, 2016).

Gestión de Averías

Es la actividad requerida para la implementación un sistema de análisis de averías y prevención vehicular, ya que una avería causa

lo siguiente: parada del vehículo lo cual conlleva a pérdida de tiempo en el que este no es utilizado para realizar su actividad; el costo de la avería; el costo de una grúa o algún otro tipo de vehículos para desplazar el vehículo averiado de ser necesario; el costo del reemplazo del vehículo; además puede ser el causante de un accidente; afectar negativamente a la imagen de la organización; y causar posibles problemas a los clientes o usuarios, si el desperfecto se produce cuando están utilizando el vehículo (Lockhart, 2016).

Control y Reducción del Consumo de Combustible

Es necesario establecer medidas para reducir y controlar el combustible usado por los vehículos, ya que esto puede generar entre el 70% y 80% de los costos variables de operación y 30% de los costos directos, y además el consumo de combustible implica la generación de emisiones contaminantes si se utiliza diésel o gasolina. Hay distintos factores involucrados en el consumo de combustible como el tipo de combustible, la estacionalidad, la forma de conducir, la antigüedad del vehículo, las condiciones del tráfico, el equipamiento del vehículo, los horarios de utilización del vehículo, las condiciones atmosféricas, la aerodinámica del vehículo, la presión de los neumáticos, el peso del vehículo, la resistencia al avance, las características del motor, el tipo de rutas, la calidad del combustible o las condiciones de la vía. Las principales medidas a tomar se deben realizar para los conductores, ya que son los que mayor impacto tienen en la reducción del consumo de combustible, se pueden tomar las siguientes acciones: brindar capacitación para conducción eficiente; el uso de dispositivos que puedan controlar y corregir la manera de conducir; identificar a los conductores que gastan menos consumos e incentivar sus buenas prácticas para nuevas estrategias; u otorgar incentivos (Lockhart, 2016).

Control y Seguimiento de los Neumáticos

Es importante realizar seguimiento y tener el control de los neumáticos ya que estos generan el 15 y el 20% de los costos variables de operación y el 3 y el 5% de los costos directos durante el ciclo de vida de un vehículo. Los neumáticos son factores a tomar en cuenta en cuanto a la seguridad de un vehículo, dependiendo del modelo y/o la marca de los neumáticos, por lo que, se deben establecer las marcas de neumáticos, y si serán premium o de bajo costo, para ser utilizados en la flota. Otros factores como la presión y el estado del neumático son importantes para la seguridad del vehículo, el consumo de combustible y el ciclo de vida útil del neumático, por lo que el estado de los neumáticos siempre debe ser verificado. El objetivo de escoger la modelo y marca del neumático es minimizar el costo por kilometraje, preservando la seguridad; para esta elección el factor más importante es el costo del neumático (Aceña, 2016).

Prevención y reducción de los accidentes de tráfico

Se establece un sistema de reducción de accidentes y para la prevención ya que un 15 y 20% de los vehículos de una flota tiene al menos un accidente en al año, lo cual representa un alto costo, debido a que el vehículo no puede ser utilizado durante un tiempo para la prestación del servicio, dependiendo de la magnitud del accidente; el costo de reparación; el costo del reemplazo del vehículo; la posible pérdida de la mercadería trasladada; y las posibles lesiones infringidas al conductor, usuarios y terceras personas. Las causas por las que se puede producir un accidente de tráfico son el fallo mecánico, condiciones del clima o las condiciones de la pista, pero la más importante es el factor humano que es causante aproximadamente del 70% de los accidentes de tráfico, teniendo como causas la falta de experiencia, conducir bajos los efectos etílicos u otras sustancias, sueño o fatiga, el exceso de

velocidad, las distracciones o no respetar las señales de tránsito (Montt, Castro & Rodríguez, 2012).

Utilización de talleres propios o externos

Se analiza si es mejor poseer talleres mecánicos propios o tercerizar esta actividad, ya que existen inversiones y costos de operación de por medio que suelen ser elevados, requiriendo de evaluaciones. Los vehículos de la flota requieren de mantenimientos y reparaciones de averías que sean realizados en un taller mecánico que cumpla con condiciones de seguridad y calidad, respecto al medioambiente y la legislación vigente. La importancia de la gestión de la flota vehículos para la organización, es el principal factor para determinar entre tener un taller mecánico propio o externo. Si la gestión de la flota soporta la principal actividad de la organización lo mejor es poseer talleres propios ya que los vehículos son reparados y se realiza el mantenimiento en un considerable tiempo reducido, para que así los vehículos se encuentren ampliamente disponibles para la prestación del servicio; se tiene ventajas como mayor control, flexibilidad y seguridad sobre los mantenimientos y reparaciones; y se genera la posibilidad de elegir que repuestos utilizar, en beneficio al ahorro de costos. Aunque si la gestión de la flota no es la actividad principal de la organización, lo mejor es que el taller a utilizar sea de un tercero, pero en consecuencia el tiempo de indisponibilidad para la prestación del servicio del vehículo es mayor, y contrario a tener un taller propio se tiene menor flexibilidad y seguridad, para los mantenimientos y reparaciones de los vehículos y no es posible la elección sobre los repuestos a emplearse (Aceña, 2016).

Costo Fijo

Son aquellos costos que son considerados necesarios e ineludibles, tales como el alquiler de algún bien, los sueldos del personal, servicios públicos, etc. Frecuentemente son constantes y no tienden a cambiar, siendo

indiferentes al grado de ventas, unidades producidas o actividades de la organización (Jiménez, 2010).

Costo Beneficio

El costo beneficio está basado en la relación existente de los beneficios en base a los costos de algún proyecto o para el estudio de si los costos operativos actuales sustentan el retorno de inversión. Entonces se debe determinar los beneficios positivos, negativos y los costos. Es necesario el impacto en cuanto al beneficio de los distintos individuos, que están implicados directa o indirectamente. Por lo que se determina un beneficio positivo cuando los costos que incurren en una inversión para una organización, en cambio un beneficio negativo cuando los costos no son justificables y hasta perjudiciales para la organización y los involucrados (Blank & Tarquin 2006).

Costos Variables de Operación

Los costos variables de operación de la flota están basados en la utilización de los vehículos, y cambian con el tiempo, por lo que son utilizados para medir como la flota es gestionada. Hay que diferenciar entre dos tipos de costos directos variables, aquellos que son obligatorios por la utilización de los vehículos, tales como son por los existentes por el mantenimiento, neumáticos, consumo de combustible, accidentes y averías; y que a su vez crecen en cuanto a mayor sea la rendimiento, el kilometraje y tiempo de actividad. Por otro lado están, aquellos de los que no dependen de la utilización de los vehículos, sino de la operatividad del servicio y de cómo se gestiona la flota (Fernández, 2016).

Nivel de Servicio

El nivel de servicio es utilizado para medir la gestión de flota ya que se busca incrementar la velocidad media, la disponibilidad y el rendimiento de los vehículos. El nivel de servicio se incrementa con la finalidad mejorar

las actividades de la gestión de la flota, incrementar la velocidad media de los vehículos, planificar mejores rutas y mejorar el servicio, capacitación de los conductores, o utilizar un software de gestión de flotas. En aquellas flotas que cuentan con rutas planificadas se considera que el nivel de servicio viene determinado por el tiempo de actividad existente entre todos los vehículos del servicio, y por lo tanto se prioriza la reducción del tiempo de actividad, pero podría suceder que se reduzca el tiempo de actividad en vehículos que tienen un menor tiempo de actividad que los demás vehículos de la flota, aumentando su velocidad media, provocando un incremento en los costos variables de operación, esto es en consecuencia al aumento del consumo de combustible para alcanzar nivel de servicio establecido (Fernández, 2016).

Formulación del problema

Problema general

¿De qué manera el desarrollo de Inteligencia de Negocios con un modelo de Machine Learning beneficia a la gestión de la flota vehicular?

Problemas específicos

- ¿De qué manera el desarrollo de Inteligencia de Negocios con un modelo de Machine Learning mejora el nivel de servicio en la gestión de la flota vehicular?
- ¿De qué manera el desarrollo de Inteligencia de Negocios con un modelo de Machine Learning mejora el costo variable de operación en la gestión de la flota vehicular?

Justificación del estudio

Justificación teórica

La presente investigación se justifica teóricamente, porque se pretende aplicar una solución de Business intelligence utilizando Machine Learning, las cuales son herramientas de gran eficiencia para solucionar la problemática planteada. Por lo que se pone en práctica los conocimientos

teóricos para dar la solución a la necesidad de la gestión de flotas vehicular en la División de Gestión de Servicios .

Justificación metodológica

La investigación ayudar a comprender como la Business intelligence, utilizando Machine Learning, ayudan a la toma decisiones de la gestión de flotas de manera efectiva; por lo que se aplicaron instrumentos de medición para la gestión de flota, para así poder determinar la validez y la confiabilidad de la solución.

Justificación tecnológica

La presente investigación tiene como finalidad el desarrollo de Business intelligence utilizando un modelo de Machine Learning, lo cual permitirá realizar la la gestión de la flota vehicular por parte de la División de Gestión de Servicios, por lo que se utilizan técnicas de Business intelligence, y ciencias de la computación, precisamente un modelo Machine Learning, que permita mostrar la información requerida de manera oportuna y con consideraciones de prever anomalías en los costos generados en la gestión de la flota vehicular.

Justificación económica

La presente investigación es justificada económicamente ya que, la solución de BI usando un modelo de machine learning, permitirá brindar información oportuna con la finalidad de reducir costos y tiempo, determinando un beneficio costo positivo para la institución, y en donde existe un alto índice de movimiento económico, por lo tanto, puede ser considerado como una inversión.

Hipótesis

Hipótesis general

El desarrollo de Inteligencia de Negocios con un modelo de Machine Learning beneficia a la gestión de la flota vehicular.

Hipótesis específicas

- El desarrollo de Inteligencia de Negocios con un modelo de Machine Learning mejora el nivel de servicio en la gestión de la flota vehicular.
- El desarrollo de Inteligencia de Negocios con un modelo de Machine Learning mejora el costo variable de operación en la gestión de la flota vehicular.

Objetivos

Objetivo general

Determinar el beneficio en la gestión de la flota vehicular mediante el desarrollo de Inteligencia de Negocios con un modelo de Machine Learning.

Objetivos específicos

- Determinar la mejora del nivel de servicio en la gestión de la flota vehicular mediante el desarrollo de Inteligencia de Negocios con un modelo de Machine Learning.
- Determinar la mejora del costo variable de operación en la gestión de la flota vehicular mediante el desarrollo de Inteligencia de Negocios con un modelo de Machine Learning.

III. METODOLOGÍA

3.1 Tipo y diseño de la investigación

La presente investigación tiene un diseño de tipo cuasi experimental, ya que se manipula la variable independiente para observar su efecto sobre la variable dependiente. Eso es porque ya se tenían los datos referentes a la gestión de flota, de manera independiente y con anterioridad a la realización de la presente investigación (Hernández, Fernández & Baptista, 2014).

El diseño de tipo cuasi experimental de un grupo con medición pre y post. Se enfoca en el análisis de los resultados de un grupo de control (antes) y un grupo de control experimental (después) para la variable dependiente así como para la variable independiente.

Esquema de diseño:

Grupo (G) O1 x O2

Donde:

X: Variable independiente y dependiente

O1: Medición antes (Pre-test)

O2 Medición después (Post-test)

3.2 Variables y operacionalización

Variables

Variable dependiente: Gestión de la flota vehicular.

Variable independiente: Inteligencia de Negocios.

Operacionalización de las variables

La operacionalización de las variables sirve para entender las definiciones, las dimensiones y los indicadores de las variables de estudio.

Matriz de Operacionalización de las variables

Tabla N°1 *Matriz de operacionalización de las variables de la investigación*

Matriz de operacionalización de las variables de la investigación

Variable	Definición Conceptual	Definición Operacional	Dimensión	Indicador	Instrumento	Escala de Medición
Gestión de la flota vehicular	Determina como es administrado un conjunto de vehículos y sus procesos interrelacionados, que tienen como principal objetivo el prestar un servicio a un cliente o servir en el apoyo de las actividades de la organización de la manera más eficiente (Fernández, 2016).	La gestión de la flota vehicular sirve para el cumplimiento y mejora del nivel de servicio y el costo de la flota vehicular, establecidos en la organización (Fernández, 2016).	Nivel de Servicio (Fernández, 2016).	Rendimiento	Ficha de Observación	Razón
				Porcentaje del costo por consumo de combustible		
Inteligencia de Negocios	Es un amplio conjunto de soluciones de TI que incluye herramientas para recolectar, analizar y reportar la información de una organización, permitiendo brindar soporte a la toma de decisiones, a través de soluciones creativas para una variedad de problemas (Maheshwari, 2015).	La Inteligencia de Negocios es para que las organizaciones ganen ventajas competitivas, teniendo el objetivo de tomar decisiones efectivas mientras se reducen los riesgos; Tiene la capacidad de ayudar a mejorar el rendimiento (Maheshwari, 2015).	Costo Variable de Operación (Fernández, 2016).	Porcentaje del costo por infracción		
				Porcentaje del costo por lavado		
				Porcentaje del costo por mantenimiento		
				Porcentaje del costo por parchado de llantas		

Tabla N°2 Indicadores

Variable	Dimensión	Indicador	Fórmula
Gestión de la flota vehicular	Nivel de Servicio (Fernández, 2016).	Rendimiento	$R = \frac{GC}{KM}$ <p>R: Rendimiento GC: Galones de combustible consumidos KM: Kilómetros recorridos</p>
	Costo Variable de Operación (Fernández, 2016).	Porcentaje del costo por consumo de combustible	$PCC = \frac{CCC}{CT} \times 100\%$ <p>PCC: Porcentaje del costo por consumo de combustible CCC: Costo por consumo de combustible CT: Costo total</p>
		Porcentaje del costo por infracción	$PI = \frac{CI}{CT} \times 100\%$ <p>PI: Porcentaje del costo por infracción CI: Costo por infracción CT: Costo total</p>
		Porcentaje del costo por lavado	$PL = \frac{CL}{CT} \times 100\%$ <p>PL: Porcentaje del costo por lavado CL: Costo por lavado CT: Costo total</p>
		Porcentaje del costo por mantenimiento	$PM = \frac{CM}{CT} \times 100\%$ <p>PM: Porcentaje del costo por mantenimiento CM: Costo por mantenimiento CT: Costo total</p>
		Porcentaje del costo por parchado de llantas	$PPL = \frac{CPL}{CT} \times 100\%$ <p>PPL: Porcentaje del costo por parchado de llantas CPL: Costo por parchado de llantas CT: Costo total</p>

3.3 Población, muestra y muestreo

Población

La población objetivo en una investigación, está definida como un conjunto finito o infinito de elementos que se están estudiando, en donde las unidades de la población poseen una característica similar la cual sirve para el estudio y determinar el origen a los datos de la investigación. Esta población es delimitada por la problemática y los objetivos del estudio (Arias, 2012).

Se consideró como población la información de la operatividad de la flota vehicular que es gestionada por la División de Gestión de Servicios, la cual se encuentra a un nivel de granularidad mensual, es decir, los meses de los que se tiene la información de consumo de combustible, mantenimiento preventivo, mantenimiento correctivo, lavado de autos, neumáticos, accidentes, tiempo de actividad, velocidad media y rendimiento de los vehículos de la flota, en un periodo de 18 meses; cabe mencionar que para todos los indicadores la población es la misma.

Muestra

La muestra, se define como un determinado grupo de individuos u objetos que son tomados de una población con la finalidad de estudiar un fenómeno estadístico. (Tamayo, 2004).

Además, si la población utilizada, es inferior a cincuenta, entonces se concluye que la muestra tiene el mismo valor que la población. (Castro, 2003).

Se debe tener en cuenta que la muestra censal es aquella porción que representa toda la población (López, 1998).

Por la cantidad de la información de la flota vehicular, que está compuesta de las distintas actividades de la operación de vehículos, y en base a los objetivos de la investigación, se decidió trabajar con el total de la población debido a que es pequeña y representa toda la muestra. Por lo que la muestra son los 18 meses en los que se generó la información de la operatividad de flota vehicular, dentro del período de Enero 2017 – Junio 2018.

3.4 Técnicas e instrumentos de recolección de datos, validez y confiabilidad

Técnicas

Observación:

Este es un método que tiene con finalidad la recolección de los datos del estudio, esto consiste en el registro sistemático, válido y confiable; de los comportamientos y de situaciones que son observables, mediante un conjunto de categorías y subcategorías (Hernández, Fernández & Baptista, 2014).

Instrumentos

Ficha de observación:

Este instrumento servirá para el registro de los datos obtenidos en las observaciones durante el proceso de recolección de datos, para así poder medir los indicadores (Valderrama, 2013).

Para recopilar la información respectiva y necesaria para el estudio, se hizo uso de la ficha de observación para todos los indicadores, para así identificar la información antes del desarrollo de la solución de Business intelligence y de la misma manera después del desarrollo.

Confiabilidad

Se determina que un instrumento medición es confiable, cuando se aplica repetidas veces al mismo individuo u objeto, y siempre se producen los mismos resultados (Hernández, Fernández & Baptista, 2014).

El coeficiente alfa es una fórmula general que permite estimar la confiabilidad de un instrumento que cuenta con ítems en los cuales se pueden asignar diferentes valores a respuestas diferentes (Aiken & Groth, 2008).

$$\alpha = \frac{k(1 - \sum s_i^2/s^2)}{k - 1}$$

Figura N°11 Coeficiente alfa

Fuente: Aiken L. & Groth G. Psychological Testing and Assessment.

Donde:

K : Número de ítems.

S_i^2 : Varianza de valores en el ítem.

S_t^2 : Varianza de valores totales del instrumento.

Validez

Para la validez, lo sustancial es comprobar si realmente el instrumento podrá medir lo que está pretendiendo medir, además de determinar su pertinencia o si es que se corresponde con los objetivos específicos y variables considerados en la investigación. Así mismo, se señala que el procedimiento puede ser realizado mediante el juicio de expertos (Arias, 2012).

3.5 Métodos de análisis de datos

En esta investigación, se tuvo en consideración que el tipo de investigación usada, fue la aplicada, así mismo el diseño utilizado fue cuasi - experimental; y por último, el método que se utilizó para analizar los datos fue el cuantitativo, ya que se obtuvieron resultados estadísticos que permitirán comprobar si la hipótesis planteada es verdadera o falsa. Para el análisis de datos cuantitativos, se debe tener en cuenta que los modelos estadísticos son representaciones que estiman como es la realidad, no representan netamente con exactitud la realidad misma; además los resultados cuantitativos siempre son interpretados según el contexto (Hernández, Fernández & Baptista, 2014).

Al ser de diseño de tipo cuasi experimental, se realizará una comparación de los resultados tomados en el pre - test y los resultados obtenidos en el post – test, después del desarrollo de la solución de Business intelligence utilizando Machine Learning. Se utilizó el entorno de desarrollo integrado RStudio utilizando el lenguaje de programación R, para realizar la estadística de los valores.

En la estadística descriptiva, se ha considerado el uso de las siguientes herramientas: coeficiente de variación, desviación estándar, media aritmética, mediana, máximo, mínimo, gráficos y tablas; para determinar los resultados obtenidos del análisis de las variables (Walpole, Myers, Myers & Ye, 2012).

Para lograr los objetivos de la investigación se consideró la prueba de Shapiro-Wilk, la cual también servirá para verificar la normalidad de los datos de estudio,

así mismo se consideró la prueba no paramétrica de Wilcoxon para verificar el efecto del desarrollo de una solución de Business intelligence utilizando un modelo de Machine Learning para la gestión de la flota vehicular en la División de Gestión de Servicios (Walpole, Myers, Myers & Ye, 2012).

3.6 Aspectos éticos

Esta investigación respeta la veracidad de los resultados y la privacidad de los datos suministrados por la División de Gestión de Servicios . Así como también se respeta la propiedad intelectual de los autores citados, reflejados en las referencias bibliográficas.

IV. RESULTADOS

4.1 Análisis Descriptivo

Para lograr los objetivos de la presente investigación se desarrolló una solución de BI usando un modelo de machine learning, para determinar cuál es su efecto en la toma de decisiones de la gestión de la flota vehicular. Se evaluaron, por parte del pre test, la información de la operatividad de la flota vehicular correspondiente a los dieciocho meses que conforma la muestra, en términos de los indicadores de porcentaje del costo por consumo de combustible, porcentaje del costo por infracción, porcentaje del costo por lavado, porcentaje del costo por mantenimiento, porcentaje del costo por parchado de llantas y rendimiento.

Como parte del post test se tomo en cuenta la información brindada por el desarrollo de la solución de business intelligence utilizando un modelo de machine learning, dicho modelo permitió realizar la predicción de los indicadores en base a los datos de los dieciocho meses que conforman la muestra. Los resultados descriptivos de estas medidas se observan en las tablas 3, 4, 5, 6, 7 y 8.

Indicadores

Tabla N°3 *Análisis descriptivo del indicador porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por consumo de combustible						
Test	Media	Mínimo	Máximo	Mediana	Desviación estándar	Coef. de variación
Pre	0.7343	0.6299	0.8345	0.7364	0.0637259	8.68%
					7	
Post	0.7334	0.6279	0.8588	0.7284	0.0694662	9.47%
					6	

Se obtuvo como media del porcentaje del costo por consumo de combustible, para el pre test un valor igual a 0.734, y para el post test el valor fue igual a 0.7334; esto significa que existe una diferencia entre el pre y post del desarrollo de la solución de BI usando un modelo de machine learning. Asimismo, el

porcentaje del costo por consumo de combustible mínimo fue de 0.6299 antes y 0.6279 después.

El coeficiente de variación del porcentaje del costo por consumo de combustible, en el pre test tuvo un resultado de 8.68% y en el post test tuvo un resultado de 9.47%, por lo que se confirma que la variabilidad de los datos no genera un cambio en gran medida, por lo que, la comparación de medias es considerada adecuada.

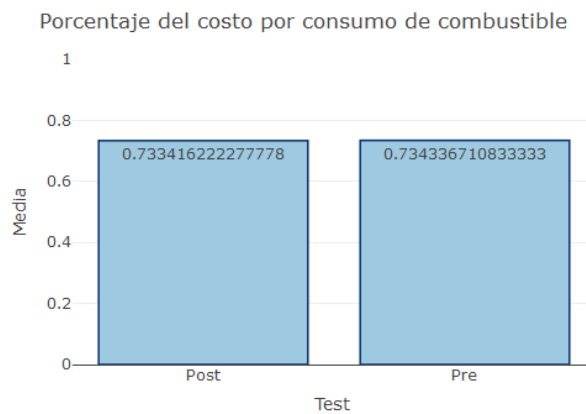


Figura N°12 Media de porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning

Fuente: Elaboración propia

Tabla N°4 *Análisis descriptivo del indicador porcentaje del costo por infracción antes de la implementación de la solución de business intelligence*

Porcentaje del costo por infracción						
Test	Media	Mínimo	Máximo	Mediana	Desviación estándar	Coef. de variación
Pre	0.002351	0.000000	0.008868	0.000975	0.00300695	127.88
	4	0	2	2		%
Pos	0.002196	0.000000	0.007340	0.001579	0.00239792	109.2%
t					6	

Se obtuvo como media del porcentaje del costo por infracción, para el pre test un valor igual a 0.0023514, y para el post test el valor fue igual a 0.002196; esto significa que existe una diferencia entre el pre y post del desarrollo de la solución de BI usando un modelo de machine learning. Asimismo, el porcentaje del costo por infracción máximo fue de 0.0088682 antes y 0.007340 después.

El coeficiente de variación del porcentaje del costo por infracción, en el pre test tuvo un resultado de 127.88% y en el post test tuvo un resultado de 109.2%, por lo que se confirma que la variabilidad de los datos no genera un cambio en gran medida, por lo que, la comparación de medias es considerada adecuada.

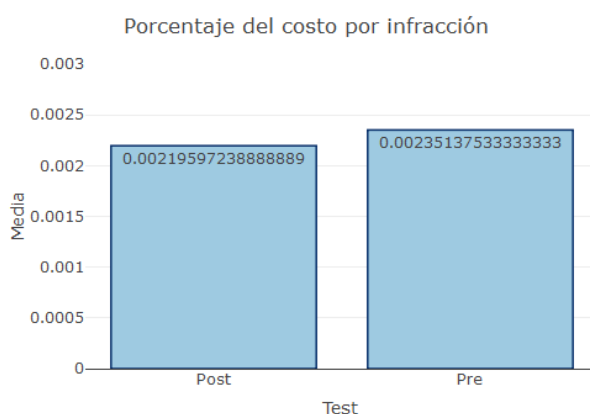


Figura N°13 Media de porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning

Fuente: Elaboración propia

Tabla N°5 *Análisis descriptivo del indicador porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por lavado						
Test	Media	Mínimo	Máximo	Mediana	Desviación estándar	Coef. de variación
Pre	0.014362	0.004213	0.025105	0.014881	0.00487789	33.96%
Post	0.014250	0.006638	0.026083	0.014317	0.004871959	34.19%

Se obtuvo como media del porcentaje del costo por lavado, para el pre test un valor igual a 0.014362, y para el post test el valor fue igual a 0.014250; esto significa que existe una diferencia entre el pre y post del desarrollo de la solución

de BI usando un modelo de machine learning. Asimismo, el porcentaje del costo por lavado mínimo fue de 0.004213 antes y 0.006638 después.

El coeficiente de variación del porcentaje del costo por lavado, en el pre test tuvo un resultado de 33.96% y en el post test tuvo un resultado de 34.19%, por lo que se confirma que la variabilidad de los datos no genera un cambio en gran medida, por lo que, la comparación de medias es considerada adecuada.

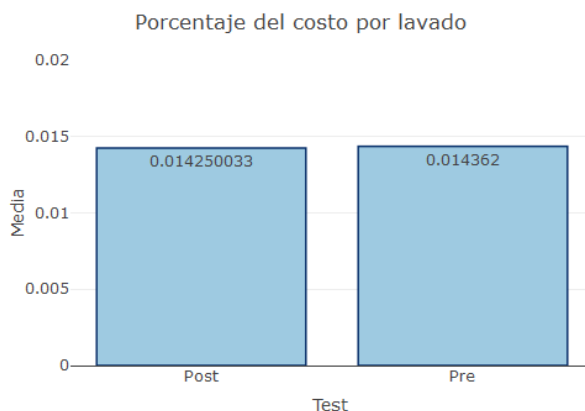


Figura N°14 Media de porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning

Fuente: Elaboración propia

Tabla N°6 *Análisis descriptivo del indicador porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por mantenimiento						
Test	Media	Mínimo	Máximo	Mediana	Desviación estándar	Coef. de variación
Pre	0.2486	0.1523	0.3541	0.2458	0.06642766	26.8%
Post	0.2479	0.1370	0.3966	0.2346	0.06818926	27.43%

Se obtuvo como media del porcentaje del costo por mantenimiento, para el pre test un valor igual a 0.2486, y para el post test el valor fue igual a 0.2486; esto significa que existe una diferencia entre el pre y post del desarrollo de la solución de BI usando un modelo de machine learning. Asimismo, el porcentaje del costo por mantenimiento mínimo fue de 0.1523 antes y 0.1370 después.

El coeficiente de variación del porcentaje del costo por mantenimiento, en el pre test tuvo un resultado de 26.8% y en el post test tuvo un resultado de 27.43%, por lo que se confirma que la variabilidad de los datos no genera un cambio en gran medida, por lo que, la comparación de medias es considerada adecuada.

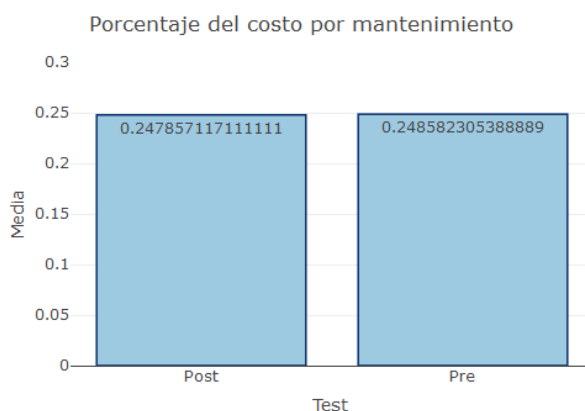


Figura N°15 Media de porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning

Fuente: Elaboración propia

Tabla N°7 *Análisis descriptivo del indicador porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por parchado de llantas						
Test	Media	Mínimo	Máximo	Mediana	Desviación estándar	Coef. de variación
Pre	0.001125	0.000038	0.002523	0.001068	0.000708434	64.83%
	2	9	4	7	6	
Post	0.001092	0.000146	0.002641	0.000994	0.000733195	65.16%
t	8	4	7	6	9	

Se obtuvo como media del porcentaje del costo por parchado de llantas, para el pre test un valor igual a 0.0011252, y para el post test el valor fue igual a 0.0010928; esto significa que existe una diferencia entre el pre y post del

desarrollo de la solución de BI usando un modelo de machine learning. Asimismo, el porcentaje del costo por parchado de llantas mínimo fue de 0.0000389 antes y 0.0001464 después.

El coeficiente de variación del porcentaje del costo por parchado de llantas, en el pre test tuvo un resultado de 64.83% y en el post test tuvo un resultado de 65.16%, por lo que se confirma que la variabilidad de los datos no genera un cambio en gran medida, por lo que, la comparación de medias es considerada adecuada.

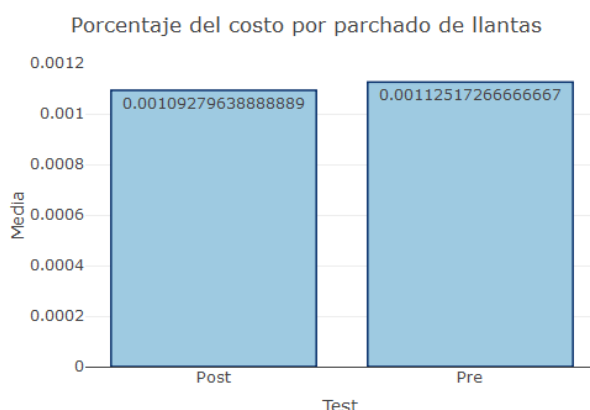


Figura N°16 Media de porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning

Fuente: Elaboración propia

Tabla N°8 *Análisis descriptivo del rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Rendimiento						
Test	Media	Mínimo	Máximo	Mediana	Desviación estándar	Coef. de variación
Pre	37.2	32.5	42.71	37.39	3.449207	9.27%
Post	37.39	32.64	44.33	37	3.858434	10.32%

Se obtuvo como media del rendimiento, para el pre test un valor igual a 37.2, y para el post test el valor fue igual a 37.39; esto significa que existe una diferencia entre el pre y post del desarrollo de la solución de BI usando un modelo de

machine learning. Asimismo, el rendimiento mínimo fue de 32.5 antes y 32.64 después.

El coeficiente de variación del rendimiento, en el pre test tuvo un resultado de 9.27% y en el post test tuvo un resultado de 10.32%, por lo que se confirma que la variabilidad de los datos no genera un cambio en gran medida, por lo que, la comparación de medias es considerada adecuada.

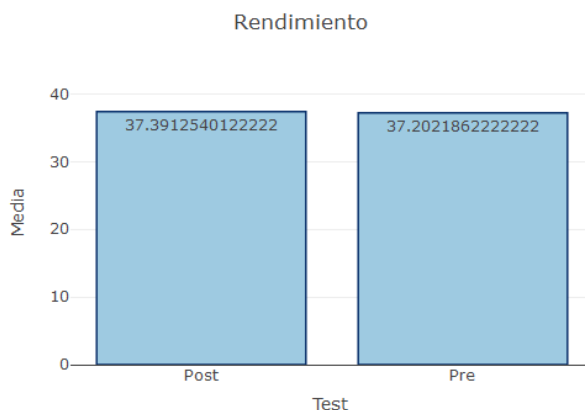


Figura N°17 Media de rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning

Fuente: Elaboración propia

4.2 Análisis Inferencial

4.2.1 Test de Shapiro Wilk

Con el propósito de escoger la prueba de hipótesis; se sometieron los datos a la validación de su distribución, refiriéndose a que si los datos tenían una distribución normal; para ello se aplicó la prueba de Shapiro-Wilk a los datos de los indicadores, dado que las muestras son menores que 50.

Indicadores:

Porcentaje del costo por consumo de combustible

Con el propósito de escoger la prueba de hipótesis; se realizó la comprobación en la distribución de los datos, refiriéndose a que si los datos de porcentaje del costo por consumo de combustible tenían una distribución normal.

Ho: Los datos presentan una distribución normal.

Ha: Los datos no presentan una distribución normal.

Tabla N°9 *Test de Shapiro Wilk de porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por consumo de combustible			
Test	Shapiro Wilk		
	Estadístico	GI	Sig.
Pre	0.95819	18	0.5671
Post	0.96427	18	0.6856

Los resultados indican que en la prueba realizada; el Sig. de porcentaje del costo por consumo de combustible antes fue de 0.5671, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el porcentaje del costo por consumo de combustible no se distribuye con normalidad.

De la misma manera, los resultados demuestran que en la prueba realizada; el Sig. de porcentaje del costo por consumo de combustible después fue de 0.6856, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el porcentaje del costo por consumo de combustible después no se distribuye con normalidad.

Este resultado sobre el Test de Shapiro Wilk también se visualiza en las Figuras 18 y 19.

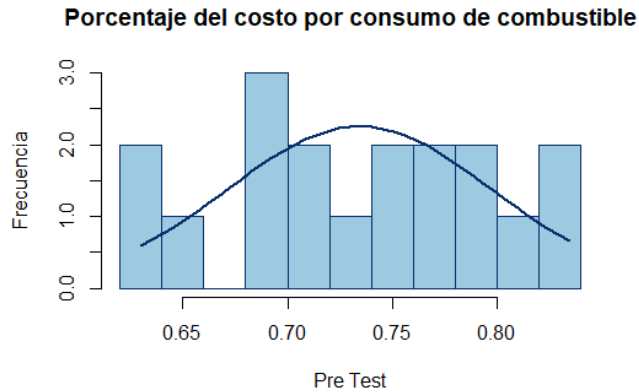


Figura N°18 Test de Shapiro Wilk de porcentaje del costo por consumo de combustible antes del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

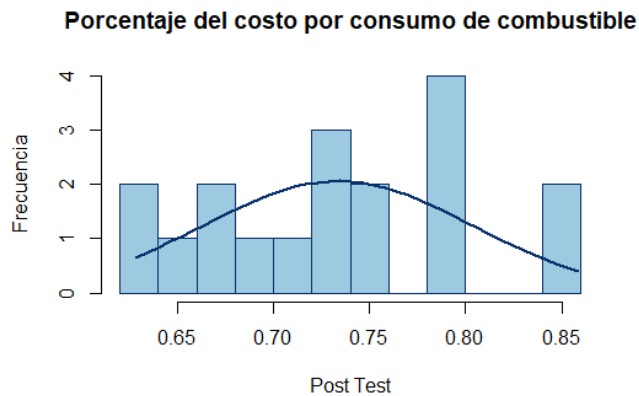


Figura N°19 Test de Shapiro Wilk de porcentaje del costo por consumo de combustible después del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Porcentaje del costo por infracción

Con el propósito de escoger la prueba de hipótesis; se sometieron los datos a la validación de su distribución, refiriéndose a que si los datos de porcentaje del costo por infracción tenían una distribución normal.

Ho: Los datos presentan una distribución normal.

Ha: Los datos no presentan una distribución normal.

Tabla N°10 Test de Shapiro Wilk de porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Test	Shapiro Wilk		
	Estadístico	Gl	Sig.
Pre	0.78613	18	0.0009739
Post	0.85036	18	0.008579

Los resultados indican que en la prueba realizada; el Sig. de porcentaje del costo por infracción antes fue de 0.0009739, cuyo valor es menor que 0.05, entonces no se rechaza la hipótesis nula, lo cual demuestra que el porcentaje del costo por infracción antes se distribuye con normalidad.

De la misma manera, los resultados demuestran que en la prueba realizada; el Sig. de porcentaje del costo por infracción después fue de 0.008579, cuyo valor es menor que 0.05, entonces no se rechaza la hipótesis nula, lo cual demuestra que el porcentaje del costo por infracción después se distribuye con normalidad.

Este resultado sobre el Test de Shapiro Wilk también se visualiza en las Figuras 20 y 21.

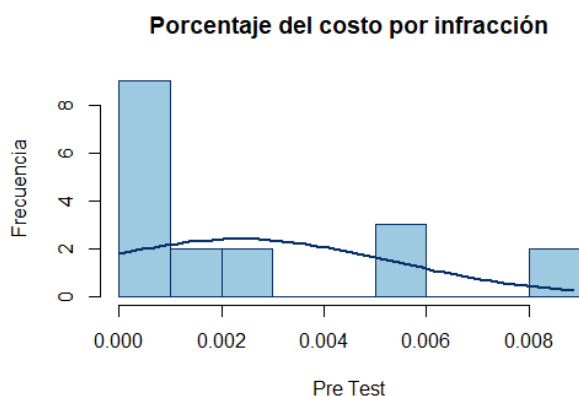


Figura N°20 Test de Shapiro Wilk de porcentaje del costo por infracción antes del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

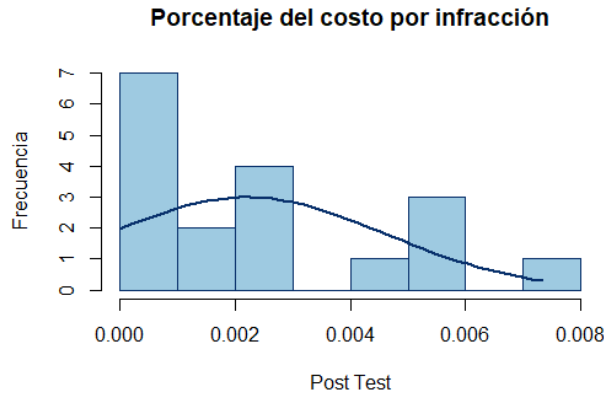


Figura N°21 Test de Shapiro Wilk de porcentaje del costo por infracción después del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Porcentaje del costo por lavado

Con el propósito de escoger la prueba de hipótesis; se sometieron los datos a la validación de su distribución, refiriéndose a que si los datos de porcentaje del costo por lavado tenían una distribución normal.

Ho: Los datos presentan una distribución normal.

Ha: Los datos no presentan una distribución normal.

Tabla N°11 *Test de Shapiro Wilk de porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por lavado			
Test	Shapiro Wilk		
	Estadístico	Gl	Sig.
Pre	0.97003	18	0.7982
Post	0.96878	18	0.7744

Los resultados indican que en la prueba realizada; el Sig. de porcentaje del costo por lavado antes fue de 0.7982, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el porcentaje del costo por lavado antes no se distribuye con normalidad.

De la misma manera, los resultados demuestran que en la prueba realizada; el Sig. de porcentaje del costo por lavado después fue de 0.7744, siendo mayor a 0.05, entonces no se rechaza la hipótesis nula, lo cual demuestra que el porcentaje del costo por lavado después no se distribuye con normalidad.

Este resultado sobre el Test de Shapiro Wilk también se visualiza en las Figuras 22 y 23.

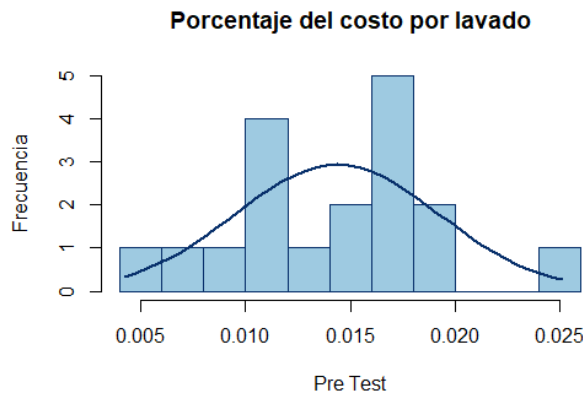


Figura N°22 Test de Shapiro Wilk de porcentaje del costo por lavado antes del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

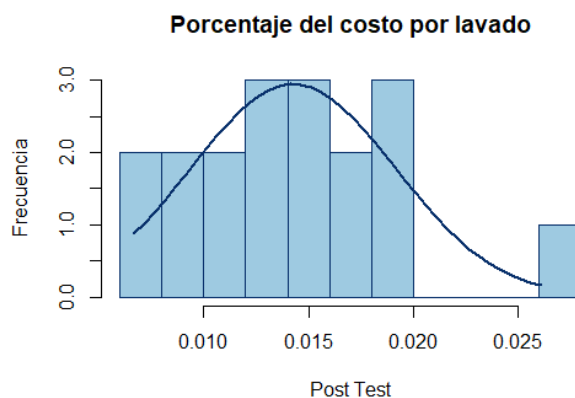


Figura N°23 Test de Shapiro Wilk de porcentaje del costo por lavado después del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Porcentaje del costo por mantenimiento

Con el propósito de escoger la prueba de hipótesis; se sometieron los datos a la validación de su distribución, refiriéndose a que si los datos de porcentaje del costo por mantenimiento tenían una distribución normal.

Ho: Los datos presentan una distribución normal.

Ha: Los datos no presentan una distribución normal.

Tabla N°12 *Test de Shapiro Wilk de porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por mantenimiento			
Test	Shapiro Wilk		
	Estadístico	GI	Sig.
Pre	0.94063	18	0.2971
Post	0.97023	18	0.802

Los resultados indican que en la prueba realizada; el Sig. de porcentaje del costo por mantenimiento antes fue de 0.2971, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el porcentaje del costo por mantenimiento antes no se distribuye con normalidad.

De la misma manera, los resultados demuestran que en la prueba realizada; el Sig. de porcentaje del costo por mantenimiento después fue de 0.802, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el porcentaje del costo por mantenimiento después no se distribuye con normalidad.

Este resultado sobre el Test de Shapiro Wilk también se visualiza en las Figuras 24 y 25.

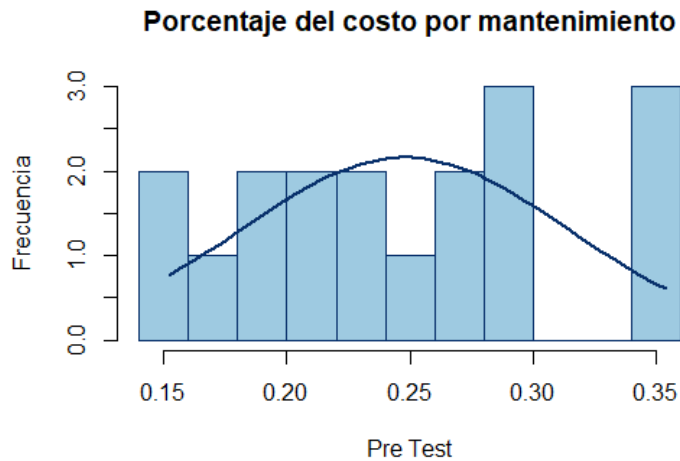


Figura N°24 Test de Shapiro Wilk de porcentaje del costo por mantenimiento antes del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

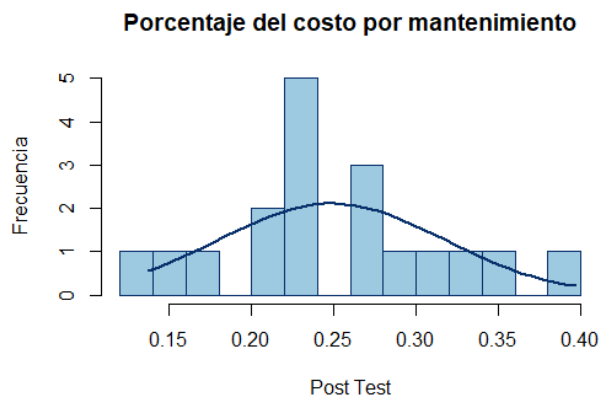


Figura N°25 Test de Shapiro Wilk de porcentaje del costo por mantenimiento después del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Porcentaje del costo por parchado de llantas

Con el propósito de escoger la prueba de hipótesis; se sometieron los datos a la validación de su distribución, refiriéndose a que si los datos de porcentaje del costo por parchado de llantas tenían una distribución normal.

Ho: Los datos presentan una distribución normal.

Ha: Los datos no presentan una distribución normal.

Tabla N°13 *Test de Shapiro Wilk de porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por parchado de llantas			
Test	Shapiro Wilk		
	Estadístico	Gl	Sig.
Pre	0.9406	18	0.2967
Post	0.93916	18	0.2804

Los resultados indican que en la prueba realizada; el Sig. de porcentaje del costo por mantenimiento antes fue de 0.2967, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el porcentaje del costo por parchado de llantas antes no se distribuye con normalidad.

De la misma manera, los resultados demuestran que en la prueba realizada; el Sig. de porcentaje del costo por parchado de llantas después fue de 0.2804, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el porcentaje del costo por parchado de llantas después no se distribuye con normalidad.

Este resultado sobre el Test de Shapiro Wilk también se visualiza en las Figuras 26 y 27.

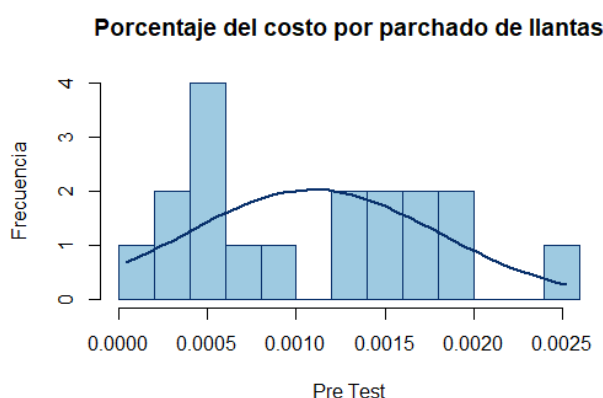


Figura N°26 Test de Shapiro Wilk de porcentaje del costo por parchado de llantas antes del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

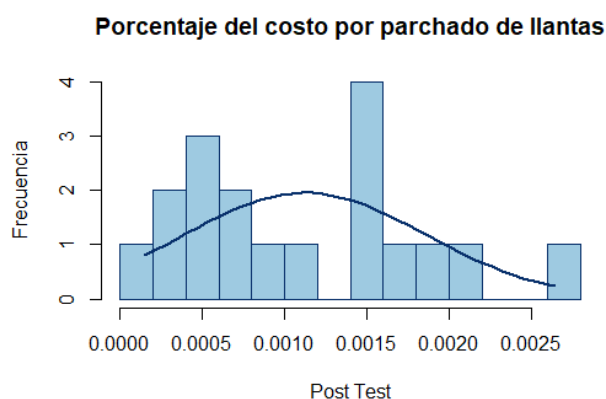


Figura N°27 Test de Shapiro Wilk de porcentaje del costo por parchado de llantas después del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Rendimiento

Con el propósito de escoger la prueba de hipótesis; se sometieron los datos a la validación de su distribución, refiriéndose a que si los datos de rendimiento tenían una distribución normal.

Ho: Los datos presentan una distribución normal.

Ha: Los datos no presentan una distribución normal.

Tabla N°14 Test de Shapiro Wilk de rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Rendimiento			
Test	Shapiro Wilk		
	Estadístico	Gl	Sig.
Pre	0.93392	18	0.2275
Post	0.91644	18	0.1118

Los resultados indican que en la prueba realizada; el Sig. de rendimiento antes fue de 0.2275, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el rendimiento antes no se distribuye con normalidad.

De la misma manera, los resultados demuestran que en la prueba realizada; el Sig. de rendimiento después fue de 0.1118, siendo mayor a 0.05, entonces la hipótesis nula es rechazada, lo cual demuestra que el rendimiento después no se distribuye con normalidad.

Este resultado sobre el Test de Shapiro Wilk también se visualiza en las Figuras 28 y 29.

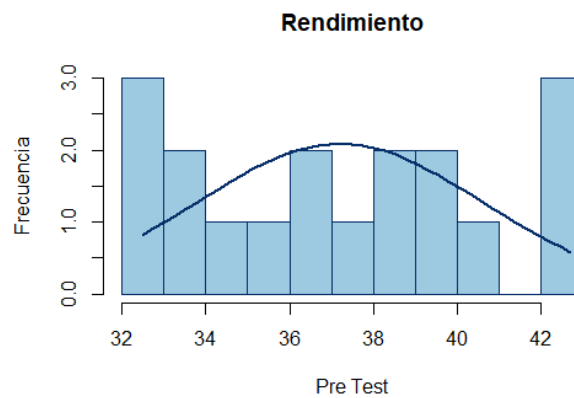


Figura N°28 Test de Shapiro Wilk rendimiento antes del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

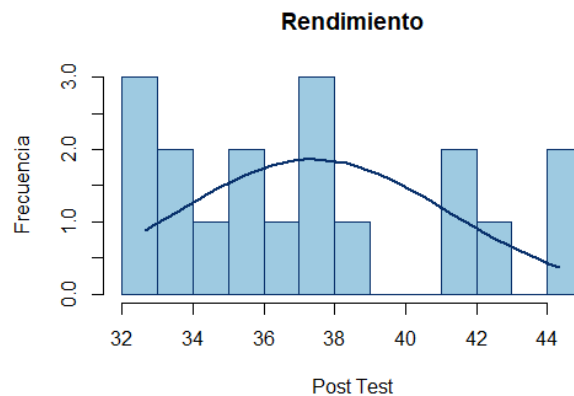


Figura N°29 Test de Shapiro Wilk rendimiento después del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

4.2.2 Prueba de Hipótesis

Como el Test de Shapiro Wilk fallo en algunas variables, los valores del post test y pre test han sido comparados mediante la prueba no paramétrica de Wilcoxon, por lo que se emparejo la información de ambos test según los meses que conforman la población, el nivel de significación que fue utilizada es de 5%.

Hipótesis de Específica N°1

H1: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el costo variable de operación en la gestión de la flota vehicular.

Indicador N°1 Porcentaje del costo por consumo de combustible

Hipótesis Estadísticas

Definición de Variables:

PCCCa: Porcentaje del costo por consumo de combustible sin la solución de BI usando un modelo de machine learning.

PCCCd: Porcentaje del costo por consumo de combustible con la solución de BI usando un modelo de machine learning.

H1o: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning no mejora el porcentaje del costo por consumo de combustible en la gestión de la flota vehicular.

$$H1o: PCCCd \geq PCCCa$$

El indicador porcentaje del costo por consumo de combustible después es mayor o igual que antes.

H1a: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el porcentaje del costo por consumo de combustible en la gestión de la flota vehicular.

$$H1a: PCCCd < PCCCa$$

El indicador porcentaje del costo por consumo de combustible después es menor que antes.

Tabla N°15 *Test de Wilcoxon para el porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por consumo de combustible			
Test	Media	Wilcoxon	
		Z	Sig. (p)
Pre	0.93392	-0.23953	0.8317
Post	0.91644		

Los resultados de la Test de Wilcoxon, definen que si p-value es mayor de 0.05, entonces la hipótesis nula es rechazada, por lo que el porcentaje del costo por consumo de combustible antes es mayor que el porcentaje del costo por consumo de combustible después de desarrollar la solución de BI usando un modelo de machine learning.

La Figura 29, muestra que el porcentaje del costo por consumo de combustible es menor en el post test (0.91644) en comparación al pre test (0.93392); por lo tanto, la solución de BI usando un modelo de machine learning mejora el porcentaje del costo por consumo de combustible.

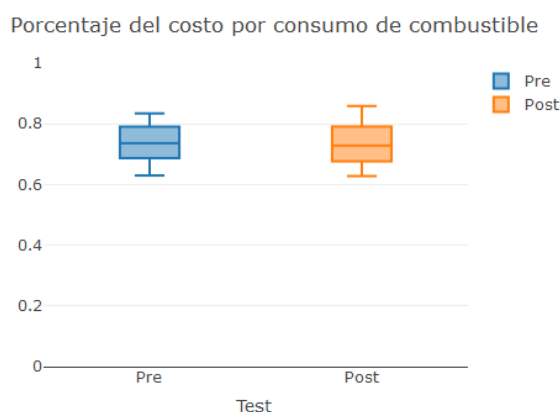


Figura N°30 Comparación del porcentaje del costo por consumo de combustible pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Indicador N°2 Porcentaje del costo por infracción
Hipótesis Estadísticas

Definición de Variables:

PCla: Porcentaje del costo por infracción sin la solución de BI usando un modelo de machine learning.

PCId: Porcentaje del costo por infracción con la solución de BI usando un modelo de machine learning.

H1o: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning no mejora el porcentaje del costo por infracción en la gestión de la flota vehicular.

$$H1o: PCId \geq PCla$$

El indicador porcentaje del costo por infracción después es mayor o igual que antes.

H1a: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el porcentaje del costo por infracción en la gestión de la flota vehicular.

$$H1a: PCId < PCla$$

El indicador porcentaje del costo por infracción después es menor que antes.

Tabla N°16 *Test de Wilcoxon para el porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por infracción			
Test	Media	Wilcoxon	
		Z	Sig. (p)
Pre	0.0023514	0.8789	0.4033
Post	0.002196		

Los resultados de la Test de Wilcoxon, definen que si p-value es mayor de 0.05, entonces la hipótesis nula es rechazada, por lo que el porcentaje del costo por infracción antes es mayor que el porcentaje del costo por infracción después de desarrollar la solución de BI usando un modelo de machine learning.

La Figura 30, muestra que el porcentaje del costo por infracción es menor en el post test (0.002196) en comparación al pre test (0.0023514); por lo tanto, la solución de BI usando un modelo de machine learning mejora el porcentaje del costo por infracción.

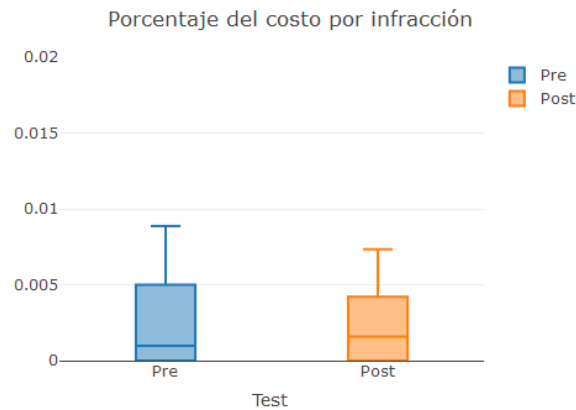


Figura N°31 Comparación del porcentaje del costo por infracción pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Indicador N°3 Porcentaje del costo por lavado

Hipótesis Estadísticas

Definición de Variables:

PCLa: Porcentaje del costo por lavado sin la solución de BI usando un modelo de machine learning.

PCLd: Porcentaje del costo por lavado con la solución de BI usando un modelo de machine learning.

H1o: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning no mejora el porcentaje del costo por lavado en la gestión de la flota vehicular.

$$H1o: PCLd \geq PCLa$$

El indicador porcentaje del costo por lavado después es mayor o igual que antes.

H1a: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el porcentaje del costo por lavado en la gestión de la flota vehicular.

$$H1a: PCLd < PCLa$$

El indicador porcentaje del costo por lavado después es menor que antes.

Tabla N°17 Test de Wilcoxon para el porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Porcentaje del costo por lavado			
Test	Media	Wilcoxon	
		Z	Sig. (p)
Pre	0.014362	0.23953	0.8317
Post	0.014250		

Los resultados de la Test de Wilcoxon, definen que si p-value es mayor de 0.05, entonces la hipótesis nula es rechazada, por lo que el porcentaje del costo por lavado antes es mayor que el porcentaje del costo por lavado después de desarrollar la solución de BI usando un modelo de machine learning.

La Figura 31, muestra que el porcentaje del costo por lavado es menor en el post test (0.014250) en comparación al pre test (0.014362); por lo tanto, la solución de BI usando un modelo de machine learning mejora el porcentaje del costo por lavado.

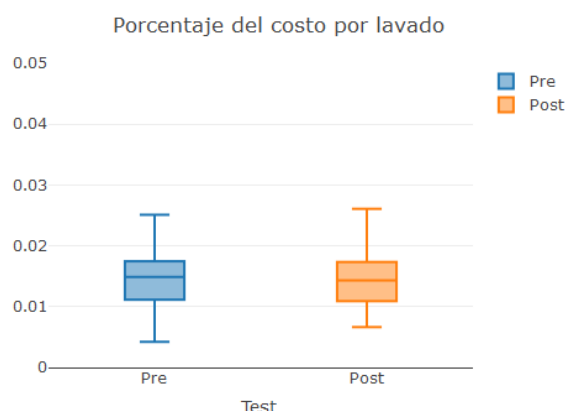


Figura N°32 Comparación del porcentaje del costo por lavado pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Indicador N°4 Porcentaje del costo por mantenimiento

Hipótesis Estadísticas

Definición de Variables:

PCMa: Porcentaje del costo por mantenimiento sin la solución de BI usando un modelo de machine learning.

PCMd: Porcentaje del costo por mantenimiento con la solución de BI usando un modelo de machine learning.

H1o: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning no mejora el porcentaje del costo por mantenimiento en la gestión de la flota vehicular.

H1o: $PCMd \geq PCMa$

El indicador porcentaje del costo por mantenimiento después es mayor o igual que antes.

H1a: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el porcentaje del costo por mantenimiento en la gestión de la flota vehicular.

H1a: $PCMd < PCMa$

El indicador porcentaje del costo por mantenimiento después es menor que antes.

Tabla N°18 *Test de Wilcoxon para el porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por mantenimiento			
Test	Media	Wilcoxon	
		Z	Sig. (p)
Pre	0.2486	-0.41373	0.7019
Post	0.2479		

Los resultados de la Test de Wilcoxon, definen que si p-value es mayor de 0.05, entonces la hipótesis nula es rechazada, por lo que el porcentaje del costo por mantenimiento antes es mayor que el porcentaje del costo por mantenimiento después de desarrollar la solución de BI usando un modelo de machine learning. La Figura 32, muestra que el porcentaje del costo por mantenimiento es menor en el post test (0.2479) en comparación al pre test (0.2486); por lo tanto, la

solución de BI usando un modelo de machine learning mejora el porcentaje del costo por mantenimiento.

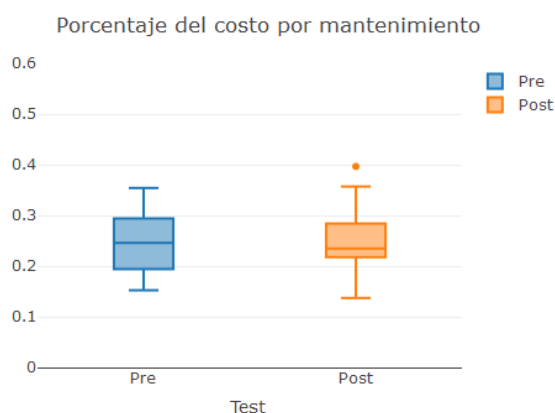


Figura N°33 Comparación del porcentaje del costo por mantenimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Indicador N°5 Porcentaje del costo por parchado de llantas

Hipótesis Estadísticas

Definición de Variables:

PCPLa: Porcentaje del costo por parchado de llantas sin la solución de BI usando un modelo de machine learning.

PCPLd: Porcentaje del costo por parchado de llantas con la solución de BI usando un modelo de machine learning.

H1o: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning no mejora el porcentaje del costo por parchado de llantas en la gestión de la flota vehicular.

$$H1o: PCPLd \geq PCPLa$$

El indicador porcentaje del costo por parchado de llantas después es mayor o igual que antes.

H1a: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el porcentaje del costo por parchado de llantas en la gestión de la flota vehicular.

H1a: PCPLd < PCPLa

El indicador porcentaje del costo por parchado de llantas después es menor que antes.

Tabla N°19 *Test de Wilcoxon para el porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning.*

Porcentaje del costo por parchado de llantas			
Test	Media	Wilcoxon	
		Z	Sig. (p)
Pre	0.0011252	-0.41373	0.7019
Post	0.0010928		

Los resultados de la Test de Wilcoxon, definen que si p-value es mayor de 0.05, entonces la hipótesis nula es rechazada, por lo que el porcentaje del costo por parchado de llantas antes es mayor que el porcentaje del costo por parchado de llantas después de desarrollar la solución de BI usando un modelo de machine learning.

La Figura 33, muestra que el porcentaje del costo por parchado de llantas es menor en el post test (0.0010928) en comparación al pre test (0.0011252); por lo tanto, la solución de BI usando un modelo de machine learning mejora el porcentaje del costo por parchado de llantas.

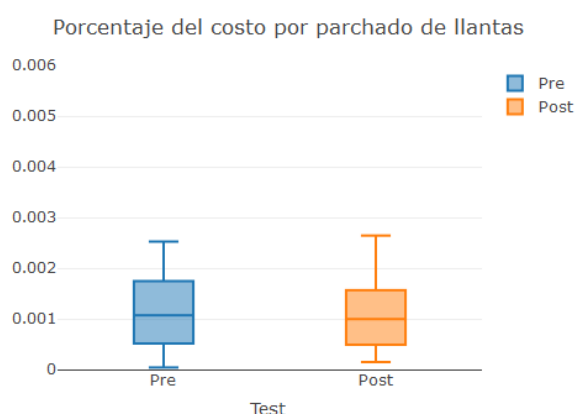


Figura N°34 Comparación del porcentaje del costo por parchado de llantas pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Por lo tanto, se puede afirmar que la hipótesis específica N°1 del presente trabajo de investigación, se ha validado, ya que se obtuvieron datos óptimos en la dimensión de costo variable de operación, luego de aplicar la solución de BI usando un modelo de machine learning.

Hipótesis de Específica N°2

H1: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el nivel del servicio en la gestión de la flota vehicular.

Indicador N°1 Rendimiento

Hipótesis Estadísticas

Definición de Variables:

Ra: Rendimiento sin la solución de BI usando un modelo de machine learning.

Rd: Rendimiento combustible con la solución de BI usando un modelo de machine learning.

H1o: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning no mejora el rendimiento en la gestión de la flota vehicular.

$$H1o: Ra \geq Rd$$

El indicador rendimiento antes es mayor o igual que después.

H1a: Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el rendimiento en la gestión de la flota vehicular.

$$H1a: Ra < Rd$$

El indicador rendimiento antes es menor que después.

Tabla N°20 Test de Wilcoxon para el rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Rendimiento			
Test	Media	Wilcoxon	
		Z	Sig. (p)
Pre	37.2	-0.84923	0.4171
Post	37.39		

Los resultados de la Test de Wilcoxon, definen que si p-value es mayor de 0.05, entonces la hipótesis nula es rechazada, por lo que el rendimiento antes es menor que el rendimiento después de desarrollar la solución de BI usando un modelo de machine learning.

La Figura 33, muestra que el rendimiento es mayor en el post test (37.39) en comparación al pre test (37.2); por lo tanto, la solución de BI usando un modelo de machine learning mejora el rendimiento.

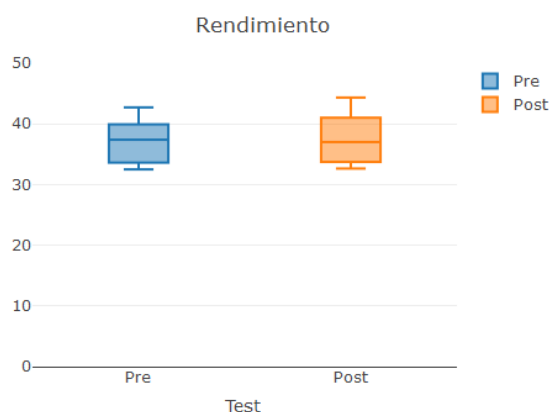


Figura N°35 Comparación del rendimiento pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

Fuente: Elaboración propia

Por lo tanto, se puede afirmar que la hipótesis específica N°2 del presente trabajo de investigación, se ha validado, ya que se obtuvieron datos óptimos en la dimensión de nivel del servicio, luego de aplicar la solución de BI usando un modelo de machine learning.

V. DISCUSIÓN

Con los resultados obtenidos en la presente investigación se analizó y se comparó el porcentaje del costo por consumo de combustible, porcentaje del costo por infracción, porcentaje del costo por lavado, porcentaje del costo por mantenimiento, porcentaje del costo por parchado de llantas y rendimiento; pre y post del desarrollo de la solución de BI usando un modelo de machine learning.

El porcentaje del costo por consumo de combustible en la medición del pre test alcanzo 0.93392 y con el desarrollo de la solución de BI usando un modelo de machine learning se redujo a 0.91644. Los resultados obtenidos indican que existe una reducción de 0.01748, con lo que se confirma que con el desarrollo de la solución de business intelligence se ha logrado una reducción de 1.87% en el porcentaje del costo por consumo de combustible.

Según la investigación realizada por Macarena, M. (2012), demostró que la solución de business intelligence redujo el porcentaje del costo por consumo de combustible en 5%, en relación a los datos obtenidos el porcentaje del costo por consumo de combustible se redujo en 1.87% en esta presente investigación.

El porcentaje del costo por infracción en la medición del pre test alcanzo 0.0023514 y con el desarrollo de la solución de BI usando un modelo de machine learning se redujo a 0.002196. Los resultados obtenidos indican que existe una reducción de 0.0001554, con lo que se confirma que con el desarrollo de la solución de business intelligence se ha logrado una reducción de 6.61% en el porcentaje del costo por infracción.

Según la investigación realizada por Laureano, O., Garza, R. y Mar, J. (2012), demostró que la solución de business intelligence redujo el porcentaje del costo por infracción en 8%, en relación a los datos obtenidos el porcentaje del costo por infracción se redujo en 6.61% en esta presente investigación.

El porcentaje del costo por lavado en la medición del pre test alcanzo 0.014362 y con el desarrollo de la solución de BI usando un modelo de machine learning se redujo a 0.014250. Los resultados obtenidos indican que existe una reducción de 0.000112, con lo que se confirma que con el desarrollo de la solución de business intelligence se ha logrado una reducción de 2% en el porcentaje del costo por lavado.

Según la investigación realizada por Laureano, O., Garza, R. y Mar, J. (2012), demostró que la solución de business intelligence redujo porcentaje del costo por lavado en 5%, en relación a los datos obtenidos el porcentaje del costo por lavado se redujo en 0.78% en esta presente investigación.

El porcentaje del costo por mantenimiento en la medición del pre test alcanzo 0.97023 y con el desarrollo de la solución de BI usando un modelo de machine learning se redujo a 0.94063. Los resultados obtenidos indican que existe una reducción de 0.0296, con lo que se confirma que con el desarrollo de la solución de business intelligence se ha logrado una reducción de 3.05% en el porcentaje del costo por mantenimiento.

Según la investigación realizada por Laureano, O., Garza, R. y Mar, J. (2012), demostró que la solución de business intelligence redujo el porcentaje del costo por mantenimiento en 6.15%, en relación a los datos obtenidos el porcentaje del costo por mantenimiento se redujo en 3.05% en esta presente investigación.

El porcentaje del costo por parchado de llantas en la medición del pre test alcanzo 0.0011252 y con el desarrollo de la solución de BI usando un modelo de machine learning se redujo a 0.0010928. Los resultados obtenidos indican que existe una reducción de 0.0000324, con lo que se confirma que con el desarrollo de la solución de business intelligence se ha logrado una reducción de 2.88% en el porcentaje del costo por parchado de llantas.

Según la investigación realizada por Laureano, O., Garza, R. y Mar, J. (2012), demostró que la solución de business intelligence redujo el porcentaje del costo por parchado de llantas en 3.31%, en relación a los datos obtenidos el porcentaje del costo por parchado de llantas se redujo en 2.88% en esta presente investigación.

El rendimiento en la medición del pre test alcanzo 37.2 y con el desarrollo de la solución de BI usando un modelo de machine learning aumentó a 37.39. Los resultados obtenidos indican que existe una reducción de 0.19, con lo que se confirma que con el desarrollo de la solución de business intelligence se ha logrado un aumento de 0.005% en el rendimiento.

Según la investigación realizada por Ricardi, M. (2013), demostró que la solución de business intelligence aumento el rendimiento en 5.45%, en relación a los datos obtenidos el rendimiento aumento en 0.005% en esta presente investigación.

VI. CONCLUSIONES

Las conclusiones de la investigación fueron las siguientes:

1. Se ha determinado el porcentaje del costo por consumo de combustible utilizando la solución de BI usando un modelo de machine learning en el proceso de la gestión de la flota vehicular, por lo que se define que sin la solución de BI usando un modelo de machine learning fue 0.93392 y con el uso de la solución de BI usando un modelo de machine learning se redujo a 0.91644. Alcanzando una reducción de 0.01748, que representa una reducción de 1.87% en el porcentaje del costo por consumo de combustible.

Así mismo se ha determinado el porcentaje del costo por infracción utilizando la solución de BI usando un modelo de machine learning en el proceso de la gestión de la flota vehicular, por lo que se define que sin la solución de BI usando un modelo de machine learning fue 0.0023514 y con el uso de la solución de BI usando un modelo de machine learning se redujo a 0.002196. Alcanzando una reducción de 0.0001554, que representa una reducción de 6.61% en el porcentaje del costo por infracción.

Además, se ha determinado el porcentaje del costo por lavado utilizando la solución de BI usando un modelo de machine learning en el proceso de la gestión de la flota vehicular, por lo que se define que sin la solución de BI usando un modelo de machine learning fue 0.014362 y con el uso de la solución de BI usando un modelo de machine learning se redujo a 0.014250. Alcanzando una reducción de 0.000112, que representa una reducción de 0.78% en el porcentaje del costo por lavado.

Así mismo se ha determinado el porcentaje del costo por mantenimiento utilizando la solución de BI usando un modelo de machine learning en el proceso de la gestión de la flota vehicular, por lo que se define que sin la solución de BI usando un modelo de machine learning fue 0.97023 y con el uso de la solución de BI usando un modelo de machine learning se redujo a 0.94063. Alcanzando una reducción de 0.0296, que representa una reducción de 3.05% en el porcentaje del costo por mantenimiento.

Además, se ha determinado el porcentaje del costo por parchado de llantas utilizando la solución de BI usando un modelo de machine learning en el proceso de la gestión de la flota vehicular, por lo que se define que sin la solución de BI

usando un modelo de machine learning fue 0.0011252 y con el uso de la solución de BI usando un modelo de machine learning se redujo a 0.0010928. Alcanzando una reducción de 0.0000324, que representa una reducción de 2.88% en el porcentaje del costo por parchado de llantas.

2. Se ha determinado el rendimiento utilizando la solución de BI usando un modelo de machine learning en el proceso de la gestión de la flota vehicular, por lo que se define que sin la solución de BI usando un modelo de machine learning fue 37.2 y con el uso de la solución de business intelligence aumento con un modelo de machine learning aumento a 37.39. Alcanzando un aumento de 0.19, que representa un aumento de 0.005% en el rendimiento.

VII. RECOMENDACIONES

Las recomendaciones para futuras investigaciones son las siguientes:

1. Se recomienda que, se debe de realizar un análisis con más características relacionadas a la operatividad de la flota vehicular y con respecto a la solución de BI usando un modelo de machine learning; con lo cual se podrá realizar análisis por los conductores de los vehículos, siniestros, compra, venta y donaciones de vehículos.
2. Considerar distintos algoritmos de regresión y clasificación para mejorar el nivel de predicción, para lo cual es necesario contar con una mayor cantidad de datos de entrenamiento, y agregar más perspectivas de análisis.
3. Evaluar y agregar nuevas características de acuerdo a la empresa donde se implemente el modelo de regresión lineal planteado en la investigación, ya que estas pueden variar.

VIII. REFERENCIAS

ACEÑA M. 2016. Gestión y control de flotas y servicios de transporte por carretera (1ra Edición). Madrid, España: Editorial CEP S.L.

ARIAS F. 2012. El Proyecto de Investigación Introducción a la metodología científica (6ta Edición). Caracas, Venezuela: Episteme.

Asset Intelligence utilises the Teradata Data Warehouse Applications platform for fleet management. 2013. Telecomworldwire (M2). Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=bwh&AN=6DUU2323202111&lang=es&site=ehost-live>

BLANK L. & TARQUIN A. 2006. Ingeniería Económica (6ta Edición). México D.F., México: Mc Graw Hill.

BRICEÑO Y. 2016. Desarrollo de un sistema informático para mejorar la gestión de mantenimiento en la empresa transportes NACIONALES S.A. Huancayo, Perú.

Business Intelligence tool. 2014. Fleet Owner, 98(1), 56. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=b9h&AN=9096488&lang=es&site=ehost-live>

CASTAÑEDA J. & GONZALES K. 2016. Plan de mejora para reducir los costos en la gestión de mantenimiento de la empresa transportes CHICLAYO S.A. Lamabayeque, Perú.

CASTRO M. 2003. El proyecto de investigación y su esquema de elaboración (2da Edición). Caracas, Venezuela: Editorial Uyapar.

CHEN C., HARDLE W. & UNWIM A. 2008. Handbook of Data Visualization (1era Edición). Alemania: Springer.

DAVENPORT T. & PATIL D. 2012. Data Scientist: The Sexiest Job of the 21st Century. Recuperado de <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

FERNÁNDEZ J. 2016. Introducción a la gestión de flotas de vehículos (1ra Edición). España: SafeCreative.

FRY B. 2008. Visualizing Data (1ra Edición). Estados Unidos: O'Reilly.

GASTALVER C. 2017. Gestión de costes y calidad del servicio de transporte por carretera (1ra Edición). España: Elearning.

GROTH G. & AIKEN L. 2008. Psychological Testing and Assessment (12va Edición). Estados Unidos: Pearson.

HERNÁNDEZ R., FERNÁNDEZ C. & BAPTISTA M. 2014. Metodología de la Investigación (6ta Edición). México: Mc Graw Hill.

IMS introduces innovative next generation fleet solution. 2013. Canada NewsWire
Recuperado de
<https://search.proquest.com/docview/1448520281?accountid=37408>

JIMÉNEZ W. 2010. Contabilidad de costos (1ra Edición). Colombia: Fundación para la Educación Superior San Mateo.

KENDALL K. & KENDALL J. 2014. Análisis y Diseño de Sistemas (9na Edición). Nueva Jersey, Estados Unidos: Pearson.

KIMBALL R. & ROSS M. 2013. The Data Warehouse Toolkit: The definitive guide to dimensional modeling (3ra Edición). Estados Unidos: Wiley.

LAUREANO O., GARZA R. & MAR J. 2012. Implementación de un sistema de gestión eficiente de flotas de transporte para la sostenibilidad económica en una empresa de transporte. México D.F., México.

LOCKHART S. 2016. Fleet: A Guide to Simplifying Vehicle Fleet Management for Small Business (1ra Edición). Canada: 90-Minute Books.

LÓPEZ, J. 1998. Proceso de Investigación (2da Edición). Caracas, Venezuela: Panapo.

MAHESHWARI A. 2014. Business Intelligence and Data Mining (1ra Edición). Estados Unidos: Business Expert Press.

MAHESHWARI A. 2015. Data Analytics Made Accessible. Iowa, Estados Unidos.

MARTÍN M. 2012. Método para la gestión eficiente del combustible en flotas de vehículos con rutas fijas. Aplicación a una empresa de construcción. Sevilla, España.

MONTT C., CASTRO F. & RODRÍGUEZ N. 2012. Análisis de Accidentes de Tránsito con Máquinas de Soporte Vectorial LS-SVM. Valparaíso, Chile.

ORACLE 2014. ¿Qué es la tecnología Java y para qué la necesito? Recuperado de https://www.java.com/es/download/faq/whatis_java.xml

ORACLE 2018. Java Platform, Standard Edition 8 Development Kit. Recuperado de <http://www.oracle.com/technetwork/java/javase/jdk-8-readme-2095712.html?ssSourceSiteId=otnes>

OZDEMIR S. 2016. Principles of Data Science (1ra Edición). Inglaterra: Packt.

PR NEWSWIRE. 2013. GE capital fleet services and MicroStrategy win the data warehousing institute's 2000 best practices in data warehousing award.

PR Newswire, pp. 1. Recuperado de <https://search.proquest.com/docview/448408587?accountid=37408>

PR NEWSWIRE. 2013. SageQuest Named NCTC Exclusive Provider for GPS Fleet Tracking and Business Intelligence. PR Newswire US. Recuperado de <http://search.ebscohost.com/login.aspx?direct=true&db=bwh&AN=201302131428PR.NEWS.USPR.NE59800&lang=es&site=ehost-live>

PR NEWSWIRE. 2015. PHH onboard telematics service delivers positive results to clients. PR Newswire, pp. 1. Recuperado de <https://search.proquest.com/docview/451419080?accountid=37408>

PROVOST F. & FAWCETT T. 2013. Data Science for Business (1ra Edición). Estados Unidos: O' Reilly Media.

PYTHON 2017. Beginner's Guide to Python. Recuperado de <https://wiki.python.org/moin/BeginnersGuide>

RICALDI M. 2013. Propuesta para la mejora de la disponibilidad de los camiones de una empresa de transportes de carga pesada, mediante el diseño de un sistema de gestión de mantenimiento. Lima, Perú.

ROBLES C. 2012. Fundamentos de administración financiera (1era Edición). Mexico: Red Tercer Milenio.

RYAN, E., L.L.B. 2013. Business intelligence changes the face of fleet management. Accountancy SA, , 45-46,48. Recuperado de <https://search.proquest.com/docview/1458788236?accountid=37408>

SANTANA J. & FARFÁN E. 2014. El arte de programar en R: un lenguaje para la estadística (1ra Edición). Mexico: Instituto Mexicano de Tecnología del Agua.

SHAI S. & SHAI B. 2014. Understanding Machine Learning: From Theory to Algorithms. Nueva York, Estados Unidos: Cambridge University Press.

TAMAYO M. 2004. El Proceso de la Investigación Científica (4ta Edición). México: Limusa.

TOBIAS S. 2014. "Right-sizing" the NS car fleet. Railway Age, 202(9), 20.
Recuperado de
<https://search.proquest.com/docview/203750709?accountid=37408>

VALDERRAMA S. 2013. Pasos para elaborar proyectos de investigación científica: cuantitativa, cualitativa y mixta (2da Edición). Perú: San Marcos.

VAN HORNE J. & WACHOWICZ J. 2010. Fundamentos de Administración Financiera (13va Edición). Estados Unidos: Pearson.

WALPOLE R., MYERS R., MYERS S. & YE K. 2012. Probabilidad y Estadística para Ingeniería y Ciencias (9na Edición). España: Pearson.

WITTEN I., FRANK E., HALL M. & PAL C. 2017. Data mining: practical machine learning tools and techniques (4ta Edición). Australia: Elsevier.

IX. ANEXOS

Anexo 1 Matriz de Contingencia

Título: Desarrollo de Inteligencia de Negocios con un modelo de Machine Learning para la gestión de la flota vehicular									
Autor: Juan Luis Vergara Pinedo									
Problema General	Objetivo General	Hipótesis General	Variable	Definición Conceptual	Definición Operacional	Dimensión	Indicador	Item	Escala
¿De qué manera el desarrollo de Inteligencia de Negocios con un modelo de Machine Learning beneficia a la toma de decisiones de la gestión de la flota vehicular ?	Determinar el beneficio de la toma de decisiones de la gestión de la flota vehicular haciendo uso de Inteligencia de Negocios con un modelo de Machine Learning.	El desarrollo de Inteligencia de Negocios con un modelo de Machine Learning beneficia a la toma de decisiones de la gestión de la flota vehicular.	V.D.: Gestión de la flota vehicular	Determina como es administrado un conjunto de vehículos y sus procesos interrelacionados, que tienen como principal objetivo el prestar un servicio a un cliente o servir en el apoyo de las actividades de la organización de la manera más eficiente (Fernández, 2016).	La gestión de la flota vehicular sirve para el cumplimiento y mejora del nivel de servicio y el costo de la flota vehicular, establecidos en la organización (Fernández, 2016).	Nivel de Servicio (Fernández, 2016).	Rendimiento	$R = \frac{GC}{KM}$ R: Rendimiento GC: Galones de combustible consumidos KM: Kilometros recorridos	Razón
							Porcentaje del costo por consumo de combustible	$PCC = \frac{CCC}{CT} \times 100\%$ PCC: Porcentaje del costo por consumo de combustible CCC: Costo por consumo de combustible CT: Costo total	Razón
¿De qué manera la Inteligencia de Negocios con un modelo de Machine Learning mejora el nivel de servicio en la toma de decisiones de la gestión de la flota vehicular ?	Determinar la mejora del nivel de servicio en la toma de decisiones de la gestión de la flota vehicular haciendo uso de Inteligencia de Negocios con un modelo de Machine Learning.	Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el nivel de servicio en la toma de decisiones de la gestión de la flota vehicular .	V.I.: Desarrollo de Inteligencia de Negocios con un modelo de Machine Learning	Es un amplio conjunto de soluciones de TI que incluye herramientas para recolectar, analizar y reportar la información de una organización, permitiendo brindar soporte a la toma de decisiones, a través de soluciones creativas para una variedad de problemas (Maheshwari, 2015).	La Inteligencia de Negocios es para que las organizaciones ganen ventajas competitivas, teniendo el objetivo de tomar decisiones efectivas mientras se reducen los riesgos; Tiene la capacidad de ayudar a mejorar el rendimiento (Maheshwari, 2015).	Costo Variable de Operación (Fernández, 2016).	Porcentaje del costo por infracción	$PI = \frac{CI}{CT} \times 100\%$ PI: Porcentaje del costo por infracción CI: Costo por infracción CT: Costo total	Razón
¿De qué manera la Inteligencia de Negocios con un modelo de Machine Learning mejora el costo variable de operación en la toma de decisiones de la gestión de la flota vehicular ?	Determinar la mejora del costo variable de operación en la toma de decisiones de la gestión de la flota vehicular haciendo uso de Inteligencia de Negocios con un modelo de Machine Learning.	Desarrollar Inteligencia de Negocios con un modelo de Machine Learning mejora el costo variable de operación en la toma de decisiones de la gestión de la flota vehicular.					Porcentaje del costo por lavado	$PL = \frac{CL}{CT} \times 100\%$ PL: Porcentaje del costo por lavado CL: Costo por lavado CT: Costo total	Razón
							Porcentaje del costo por mantenimiento	$PM = \frac{CM}{CT} \times 100\%$ PM: Porcentaje del costo por mantenimiento CM: Costo por mantenimiento CT: Costo total	Razón
							Porcentaje del costo por parchado de llantas	$PPL = \frac{CPL}{CT} \times 100\%$ PPL: Porcentaje del costo por parchado de llantas CPL: Costo por parchado de llantas CT: Costo total	Razón

Anexo 2 Fichas de Observación

FICHA DE OBSERVACIÓN

Nº FICHA DE OBSERVACIÓN:	1
VALIDADOR:	Vergara Pinedo Juan Luis
POBLACIÓN:	Información mensual de la operatividad de la flota vehicular
PERIODO:	Enero 2017 - Junio 2018
DIMENSIÓN:	Costo Variable de Operación
INDICADOR:	Porcentaje del costo por consumo de combustible
FÓRMULA:	Costo por consumo de combustible/Costo Total

ITEM	MES	PRE - TEST	RESULTADO PRE - TEST	POST - TEST	RESULTADO POST - TEST
1	Ene-17	223881.22	0.790694507	222917.1474	0.787289635
		283145.03		283145.03	
2	Feb-17	191539.15	0.824394569	196497.6852	0.845736366
		232339.17		232339.17	
3	Mar-17	222293.95	0.762379182	219930.497	0.754273485
		291579.25		291579.25	
4	Abr-17	203436.43	0.687034399	208871.6346	0.705389875
		296108.07		296108.07	
5	May-17	225572.36	0.771793935	231205.5141	0.791067724
		292270.19		292270.19	
6	Jun-17	219357.89	0.69641626	213079.5284	0.676483751
		314981		314981	
7	Jul-17	201102.87	0.629884969	200481.8405	0.627939809
		319269.2		319269.2	
8	Ago-17	219687.14	0.682507441	222289.1058	0.690591032
		321882.41		321882.41	
9	Set-17	198552.17	0.753425945	189872.3616	0.720489549
		263532.43		263532.43	
10	Oct-17	216712	0.80929538	213652.9079	0.79787142
		267778.62		267778.62	
11	Nov-17	213476.69	0.720565702	214912.4102	0.725411809
		296262.63		296262.63	
12	Dic-17	206018.95	0.791614051	207255.7964	0.796366551
		260251.76		260251.76	
13	Ene-18	204483.42	0.83446946	210440.7592	0.858780564
		245046.02		245046.02	
14	Feb-18	201949.25	0.713145278	183270.0496	0.647183243
		283181.08		283181.08	
15	Mar-18	200322.4	0.717999941	204061.85	0.73140296
		279000.58		279000.58	
16	Abr-18	196938.26	0.638346053	195431.466	0.633462004
		308513.32		308513.32	
17	May-18	226509.56	0.752263187	223725.8453	0.743018164
		301104.14		301104.14	
18	Jun-18	164977.7	0.641830536	171893.0478	0.66873406
		257042.46		257042.46	

FICHA DE OBSERVACIÓN

N° FICHA DE OBSERVACIÓN:	2
VALIDADOR:	Vergara Pinedo Juan Luis
POBLACIÓN:	Información mensual de la operatividad de la flota vehicular
PERIODO:	Enero 2017 - Junio 2018
DIMENSIÓN:	Costo Variable de Operación
INDICADOR:	Porcentaje del costo por infracción
FÓRMULA:	Costo por infracción/Costo Total

ITEM	MES	PRE - TEST	RESULTADO PRE - TEST	POST - TEST	RESULTADO POST - TEST
1	Ene-17	2511	0.008868247	2078.394343	0.007340388
		283145.03		283145.03	
2	Feb-17	0	0	0	0
		232339.17		232339.17	
3	Mar-17	1458	0.005000356	1482.208889	0.005083383
		291579.25		291579.25	
4	Abr-17	810	0.002735488	596.1854545	0.002013405
		296108.07		296108.07	
5	May-17	0	0	0	0
		292270.19		292270.19	
6	Jun-17	0	0	0	0
		314981		314981	
7	Jul-17	0	0	0	0
		319269.2		319269.2	
8	Ago-17	1782	0.005536183	1780.301616	0.005530907
		321882.41		321882.41	
9	Set-17	324	0.00122945	298.0927273	0.001131142
		263532.43		263532.43	
10	Oct-17	1566.94	0.005851625	1482.208889	0.005535203
		267778.62		267778.62	
11	Nov-17	729	0.002460655	741.1044444	0.002501512
		296262.63		296262.63	
12	Dic-17	324	0.001244948	298.0927273	0.001145401
		260251.76		260251.76	
13	Ene-18	0	0	0	0
		245046.02		245046.02	
14	Feb-18	2324	0.008206763	1192.370909	0.004210631
		283181.08		283181.08	
15	Mar-18	131.14	0.000470035	596.1854545	0.002136861
		279000.58		279000.58	
16	Abr-18	222.44	0.000721006	894.2781818	0.00289867
		308513.32		308513.32	
17	May-18	0	0	0	0
		301104.14		301104.14	
18	Jun-18	0	0	0	0
		257042.46		257042.46	

FICHA DE OBSERVACIÓN

N° FICHA DE OBSERVACIÓN:	3
VALIDADOR:	Vergara Pinedo Juan Luis
POBLACIÓN:	Información mensual de la operatividad de la flota vehicular
PERIODO:	Enero 2017 - Junio 2018
DIMENSIÓN:	Costo Variable de Operación
INDICADOR:	Porcentaje del costo por lavado
FÓRMULA:	Costo por lavado/Costo Total

ITEM	MES	PRE - TEST	RESULTADO PRE - TEST	POST - TEST	RESULTADO POST - TEST
1	Ene-17	4015	0.014180012	4157.380238	0.014682865
		283145.03		283145.03	
2	Feb-17	4325	0.018615027	4555.592216	0.019607508
		232339.17		232339.17	
3	Mar-17	7320	0.025104667	7605.341926	0.026083276
		291579.25		291579.25	
4	Abr-17	3495	0.011803123	3642.382162	0.012300854
		296108.07		296108.07	
5	May-17	3930	0.013446462	3767.457913	0.012890326
		292270.19		292270.19	
6	Jun-17	5375	0.017064521	4394.263188	0.013950883
		314981		314981	
7	Jul-17	4975	0.015582461	5013.543828	0.015703187
		319269.2		319269.2	
8	Ago-17	5620	0.017459792	5582.640861	0.017343728
		321882.41		321882.41	
9	Set-17	2555	0.009695201	2586.754346	0.009815696
		263532.43		263532.43	
10	Oct-17	5000	0.01867214	5005.465118	0.01869255
		267778.62		267778.62	
11	Nov-17	4950	0.016708148	4950.498348	0.016709831
		296262.63		296262.63	
12	Dic-17	2900	0.011143056	2837.805965	0.01090408
		260251.76		260251.76	
13	Ene-18	2870	0.011712086	2680.032891	0.010936855
		245046.02		245046.02	
14	Feb-18	4885	0.017250446	4399.965803	0.01553764
		283181.08		283181.08	
15	Mar-18	2910	0.010430086	2646.663725	0.00948623
		279000.58		279000.58	
16	Abr-18	2385	0.007730622	2047.867979	0.006637859
		308513.32		308513.32	
17	May-18	5331	0.017704838	5471.198392	0.018170452
		301104.14		301104.14	
18	Jun-18	1083	0.004213312	1811.32018	0.007046774
		257042.46		257042.46	

FICHA DE OBSERVACIÓN

N° FICHA DE OBSERVACIÓN:	4
VALIDADOR:	Vergara Pinedo Juan Luis
POBLACIÓN:	Información mensual de la operatividad de la flota vehicular
PERIODO:	Enero 2017 - Junio 2018
DIMENSIÓN:	Costo Variable de Operación
INDICADOR:	Porcentaje del costo por mantenimiento
FÓRMULA:	Costo por mantenimiento/Costo Total

ITEM	MES	PRE - TEST	RESULTADO PRE - TEST	POST - TEST	RESULTADO POST - TEST
1	Ene-17	52577.81	0.185692152	56772.5914	0.200507109
		283145.03		283145.03	
2	Feb-17	36070.02	0.155247262	37449.9729	0.161186652
		232339.17		232339.17	
3	Mar-17	60107.3	0.206143956	63454.09203	0.217622111
		291579.25		291579.25	
4	Abr-17	87846.64	0.296670874	69654.45503	0.235233221
		296108.07		296108.07	
5	May-17	62682.83	0.214468776	66127.83063	0.226255817
		292270.19		292270.19	
6	Jun-17	90148.11	0.286201739	86401.42949	0.274306798
		314981		314981	
7	Jul-17	113041.33	0.354062747	113972.7402	0.356980066
		319269.2		319269.2	
8	Ago-17	94628.27	0.293983974	97194.53215	0.301956644
		321882.41		321882.41	
9	Set-17	61436.26	0.233125995	58544.75134	0.222153878
		263532.43		263532.43	
10	Oct-17	44149.68	0.164873805	40687.32718	0.151943897
		267778.62		267778.62	
11	Nov-17	76556.94	0.258409034	78409.18825	0.264661082
		296262.63		296262.63	
12	Dic-17	50498.81	0.194038304	58607.19922	0.225194247
		260251.76		260251.76	
13	Ene-18	37332.6	0.152349342	33575.92071	0.137018837
		245046.02		245046.02	
14	Feb-18	73822.83	0.260691251	74588.17123	0.263393908
		283181.08		283181.08	
15	Mar-18	75237.04	0.269666249	79211.89834	0.283913024
		279000.58		279000.58	
16	Abr-18	108807.62	0.352683703	122359.9827	0.396611669
		308513.32		308513.32	
17	May-18	69013.58	0.229201697	70466.01619	0.234025398
		301104.14		301104.14	
18	Jun-18	90971.76	0.353917248	82643.55643	0.321517139
		257042.46		257042.46	

FICHA DE OBSERVACIÓN

N° FICHA DE OBSERVACIÓN:	5
VALIDADOR:	Vergara Pinedo Juan Luis
POBLACIÓN:	Información mensual de la operatividad de la flota vehicular
PERIODO:	Enero 2017 - Junio 2018
DIMENSIÓN:	Costo Variable de Operación
INDICADOR:	Porcentaje del costo por parchado de llantas
FÓRMULA:	Costo por parchado de llantas/Costo Total

ITEM	MES	PRE - TEST	RESULTADO PRE - TEST	POST - TEST	RESULTADO POST - TEST
1	Ene-17	160	0.000565081	153	0.000540359
		283145.03		283145.03	
2	Feb-17	405	0.001743141	413.9444444	0.001781639
		232339.17		232339.17	
3	Mar-17	400	0.00137184	330.8888889	0.001134816
		291579.25		291579.25	
4	Abr-17	520	0.001756116	629.8333333	0.002127039
		296108.07		296108.07	
5	May-17	85	0.000290827	142.6666667	0.000488133
		292270.19		292270.19	
6	Jun-17	100	0.000317479	46.11111111	0.000146393
		314981		314981	
7	Jul-17	150	0.000469823	137.5	0.000430671
		319269.2		319269.2	
8	Ago-17	165	0.00051261	275	0.000854349
		321882.41		321882.41	
9	Set-17	665	0.002523409	525.4444444	0.001993851
		263532.43		263532.43	
10	Oct-17	350	0.00130705	412.5	0.001540452
		267778.62		267778.62	
11	Nov-17	550	0.001856461	463.4444444	0.001564303
		296262.63		296262.63	
12	Dic-17	510	0.001959641	687.5	0.002641673
		260251.76		260251.76	
13	Ene-18	360	0.001469112	367.2222222	0.001498585
		245046.02		245046.02	
14	Feb-18	200	0.000706262	203.3333333	0.000718033
		283181.08		283181.08	
15	Mar-18	400	0.001433689	417.3333333	0.001495815
		279000.58		279000.58	
16	Abr-18	160	0.000518616	86.83333333	0.000281457
		308513.32		308513.32	
17	May-18	250	0.000830278	222.2222222	0.000738024
		301104.14		301104.14	
18	Jun-18	10	0.000038904	71.33333333	0.000038904
		257042.46		257042.46	

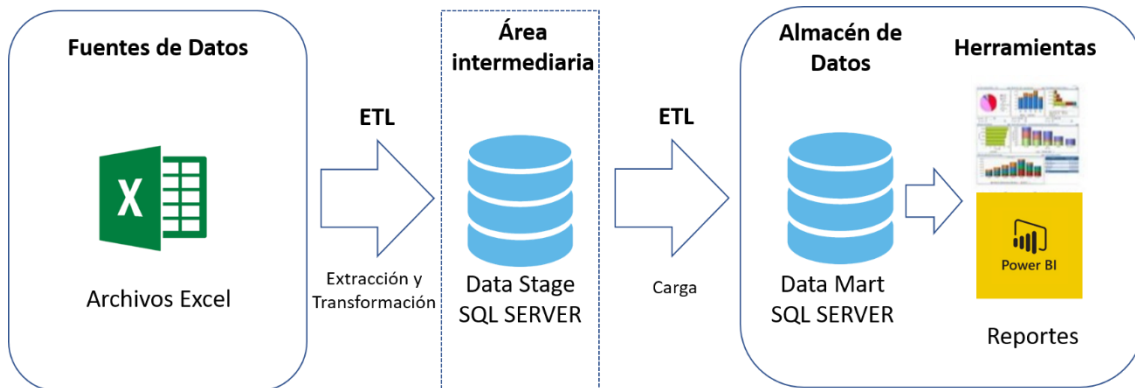
FICHA DE OBSERVACIÓN

N° FICHA DE OBSERVACIÓN:	6
VALIDADOR:	Vergara Pinedo Juan Luis
POBLACIÓN:	Información mensual de la operatividad de la flota vehicular
PERIODO:	Enero 2017 - Junio 2018
DIMENSIÓN:	Nivel de Servicio
INDICADOR:	Rendimiento
FÓRMULA:	Kilometros Recorridos/Galones Consumidos

ITEM	MES	PRE - TEST	RESULTADO PRE - TEST	POST - TEST	RESULTADO POST - TEST
1	Ene-17	1298.28325	32.9342	1298.28325	32.63724911
		47.9867		47.9867	
2	Feb-17	1162.099222	32.528678	1162.099222	32.9832227
		42.786502		42.786502	
3	Mar-17	1380.052564	32.497923	1380.052564	33.72758253
		49.941846		49.941846	
4	Abr-17	1367.9104	34.951466	1367.9104	34.93187326
		48.01472		48.01472	
5	May-17	1376.923248	33.4578	1376.923248	32.91209161
		51.807544		51.807544	
6	Jun-17	1362.132461	33.612309	1362.132461	33.41268117
		50.30439		50.30439	
7	Jul-17	1331.092354	42.122405	1331.092354	42.6306475
		47.737012		47.737012	
8	Ago-17	1370.359045	42.71402	1370.359045	44.16680561
		52.3651		52.3651	
9	Set-17	1388.768675	42.3244	1388.768675	44.33282111
		45.586775		45.586775	
10	Oct-17	1288.396034	38.196428	1288.396034	35.92477944
		49.782586		49.782586	
11	Nov-17	1372.812266	39.918933	1372.812266	41.01191556
		52.17864		52.17864	
12	Dic-17	1262.125655	38.125321	1262.125655	37.9134714
		48.193264		48.193264	
13	Ene-18	1257.210389	37.826051	1257.210389	36.6682923
		47.469038		47.469038	
14	Feb-18	1269.689074	35.423544	1269.689074	35.65438686
		46.699523		46.699523	
15	Mar-18	1357.679329	39.632737	1357.679329	38.08763472
		49.204525		49.204525	
16	Abr-18	1380.173629	40.012231	1380.173629	41.20904221
		47.072768		47.072768	
17	May-18	1349.318414	36.945217	1349.318414	37.3328401
		50.229232		50.229232	
18	Jun-18	1170.304105	36.415689	1170.304105	37.50523503
		41.48346		41.48346	

Anexo 3 Desarrollo del Proyecto

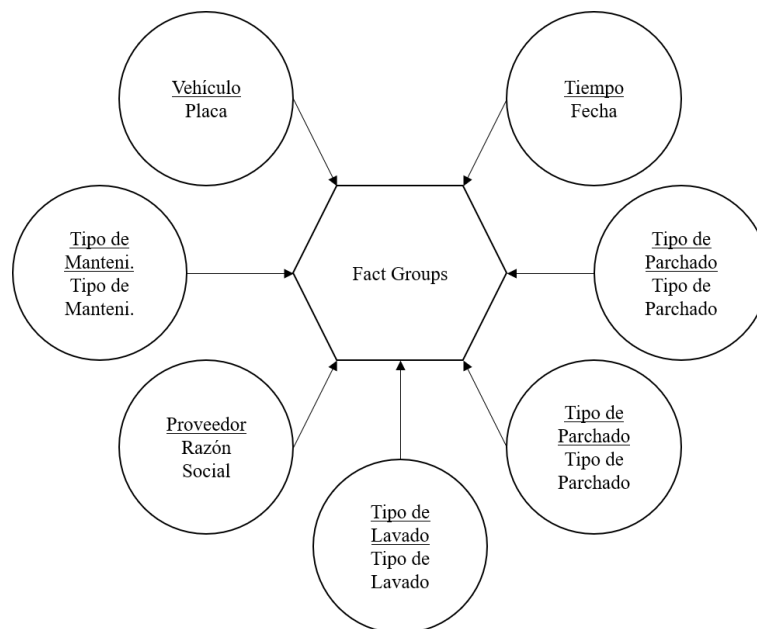
Diseño de la Arquitectura



Modelamiento Dimensional

La siguiente sección presenta las dimensiones y áreas de análisis que serán incluidas en este proyecto.

Dimensiones



Nombre	Descripción
Proveedor	Información de los distintos proveedores que ofrecen servicios para las distintas actividades de la flota vehicular.

Tiempo	Información asociada con la fecha en que sucedió una actividad en específica.
Tipo de Combustible	Información del tipo de combustible consumido por los vehículos.
Tipo de Lavado	Información del tipo de lavado que se realizan a los vehículos.
Tipo de Mantenimiento	Información del tipo de mantenimiento vehicular realizado.
Tipo de Parchado de Llanta	Información del tipo de parchado de llantas que se realizan a los vehículos.
Vehículo	Información asociada a cada vehículo perteneciente a la flota vehicular.

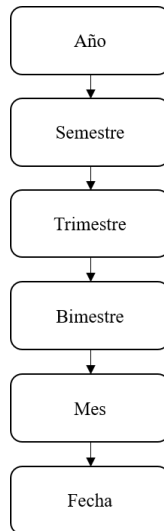
Detalle de las Dimensiones

Dimensión Proveedor

Razón Social

Atributo	Descripción	Posibles Valores
Razón Social	Razón social del proveedor.	Cualquier proveedor que brinde servicios para la flota vehicular.

Dimensión Tiempo



Atributo	Descripción	Posibles Valores
Año	Año de análisis.	2017, 2018.
Semestre	Semestre de análisis.	Primer Semestre, Segundo Semestre.
Trimestre	Trimestre de análisis.	Primer Trimestre, Segundo Trimestre, Tercer Trimestre, Cuarto Trimestre.
Bimestre	Bimestre de análisis.	Primer Bimestre, Segundo Bimestre, Tercer Bimestre, Cuarto Bimestre, Quinto Bimestre, Sexto Bimestre.
Mes	Mes de análisis.	Enero, Febrero, Marzo, Abril, Mayo, Junio, Julio, Agosto, Septiembre, Octubre, Noviembre, Diciembre.
Día	Día del mes de análisis.	1, 2, 3, 4, ... 30 y/o 31.

Dimensión Tipo de Combustible

Tipo de Combustible

Atributo	Descripción	Posibles Valores
Tipo de combustible	Es el tipo de combustible.	Diesel D2, Gasohol 84, Gasohol 90, Gasohol 95, Gasohol 97, Gasohol 98, Gas Natural Vehicular.

Dimensión Tipo de Lavado

Tipo de Lavado

Atributo	Descripción	Posibles Valores
Tipo de lavado	Es el tipo de lavado.	Simple, Salón.

Dimensión Tipo de Mantenimiento

Tipo de Mantenimiento

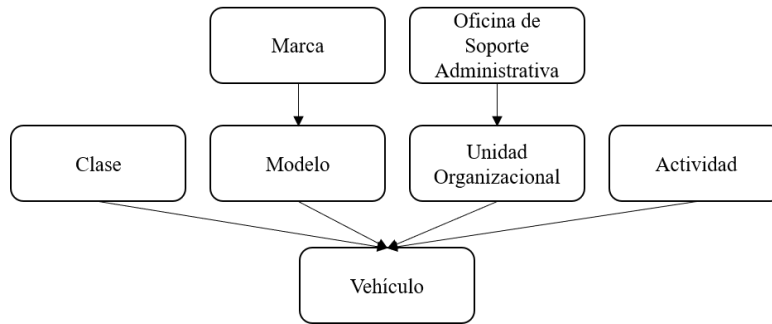
Atributo	Descripción	Posibles Valores
Tipo de mantenimiento	Es el tipo de mantenimiento.	Correctivo, Preventivo.

Dimensión Tipo de Parchado de Llanta

Tipo de Parchado de Llanta

Atributo	Descripción	Posibles Valores
Tipo de parchado de llanta	Es el tipo de parchado de llanta.	Parchado, Enllante, Rotación.

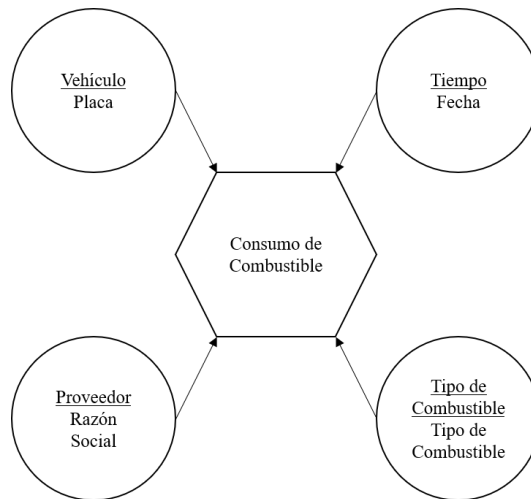
Dimensión Vehículo



Atributo	Descripción	Posibles Valores
Vehículo.	Vehículo del análisis.	Distintas placas de los vehículos que forman parte de la flota vehicular.
Marca	Marca del vehículo	Marca de cada vehículo.
Modelo	Modelo del vehículo	Modelo de cada vehículo.
Clase	Clase del vehículo	Clase de cada vehículo
Actividad	Actividad o estado actual del vehículo	Activo, Inactivo, Afectación en uso, De baja.
Unidad Organizacional	Unidad Organizacional a la que pertenece el vehículo	Unidad Organizacional de cada vehículo.
Oficina de Soporte Administrativa	Oficina de Soporte Administrativa a la que está asignada el vehículo	Oficina de Soporte Administrativa de cada vehículo.

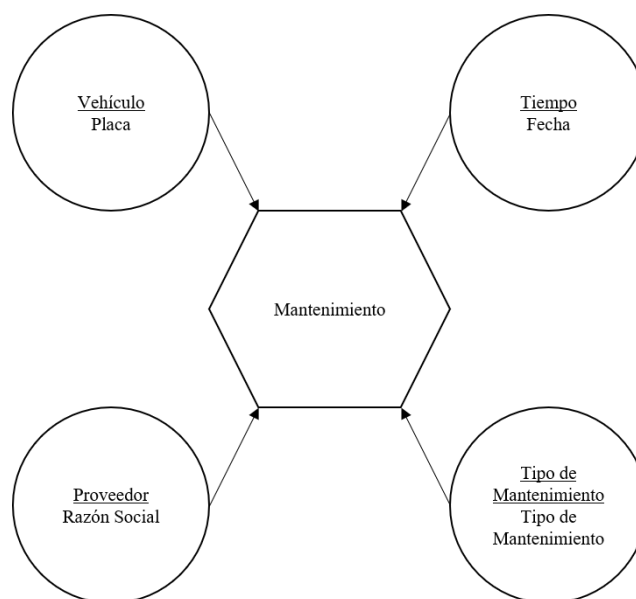
Tablas de Hechos

Consumo de combustible



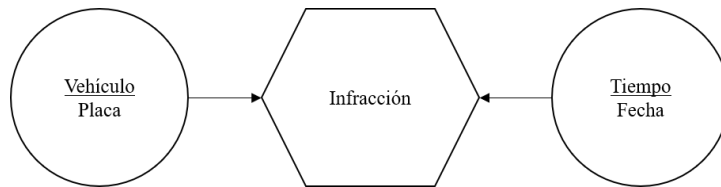
Métrica	Descripción	Regla de agregación por defecto
Costo de combustible	Es el costo de un tipo combustible brindado por un proveedor y utilizado por los vehículos en un periodo de tiempo.	Sum
Porcentaje de costo de combustible	Porcentaje del costo de combustible en base al costo total realizado por la flota vehicular.	N/A

Mantenimiento



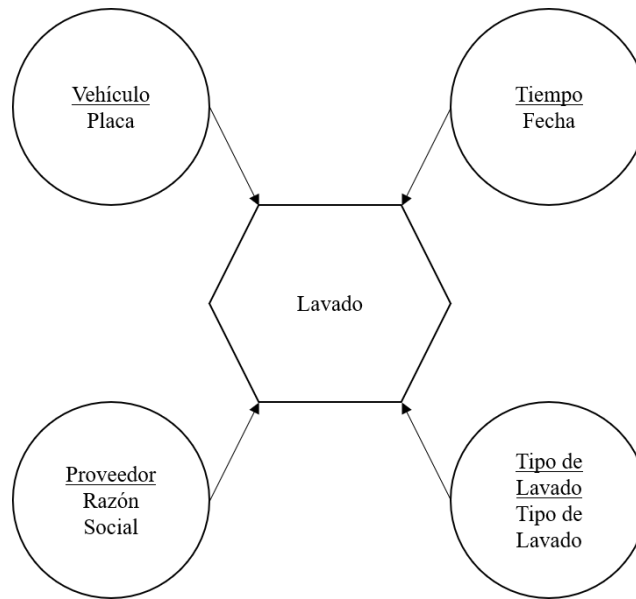
Métrica	Descripción	Regla de agregación por defecto
Costo de mantenimiento	Es el costo de un tipo mantenimiento brindado por un proveedor y realizado a los vehículos en un periodo de tiempo.	Sum
Porcentaje de costo de mantenimiento	Porcentaje del costo de mantenimiento en base al costo total realizado por la flota vehicular.	N/A

Infracción



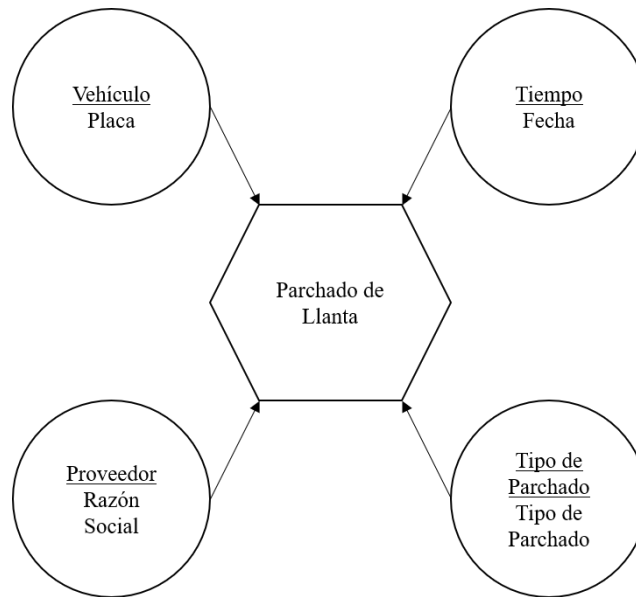
Métrica	Descripción	Regla de agregación por defecto
Costo de infracción	Es el costo de infracción realizado por los vehículos en un periodo de tiempo.	Sum
Porcentaje de costo de mantenimiento	Porcentaje del costo de infracción en base al costo total realizado por la flota vehicular.	N/A

Lavado



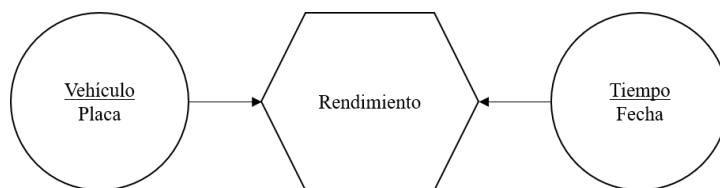
Métrica	Descripción	Regla de agregación por defecto
Costo de lavado	Es el costo de un tipo de lavado brindado por un proveedor y realizado a los vehículos en un periodo de tiempo.	Sum
Porcentaje de costo de lavado	Porcentaje del costo de lavado en base al costo total realizado por la flota vehicular.	N/A

Parchado de Llanta



Métrica	Descripción	Regla de agregación por defecto
Costo de parchado	Es el costo de un tipo de parchado de llanta brindado por un proveedor y realizado a los vehículos en un periodo de tiempo.	Sum
Porcentaje de costo de parchado	Porcentaje del costo de parchado de llanta en base al costo total realizado por la flota vehicular.	N/A

Rendimiento



Métrica	Descripción	Regla de agregación por defecto
Kilómetro Recorrido	Es la cantidad de kilometro recorridos por un vehículo en un periodo determinado.	Sum

Galones de Combustible Consumidos	Es la cantidad de galones de combustible consumido por un vehículo en un periodo determinado.	Sum
Rendimiento Vehicular	Es el rendimiento de un vehículo en un periodo de tiempo.	N/A

Base de Datos Intermedia – Data Stage

Stg_ConsumoCombustible
ConsumoCombustible_SKey
Vehiculo_Placa
Vehiculo_SKey
TipoCombustible_Codigo
TipoCombustible_SKey
Proveedor_Codigo
Proveedor_SKey
Tiempo_SKey
ConsumoCombustible_Fecha
ConsumoCombustible_Costo

Stg_Vehiculo
Vehiculo_SKey
Vehiculo_Placa
Vehiculo_Marca
Vehiculo_Modelo
Vehiculo_Clase
Vehiculo_UnidadOrganizaci...
Vehiculo_OficinaSoporteAd...
Vehiculo_Actividad

Stg_TipoMantenimiento
TipoMantenimiento_SKey
TipoMantenimiento_Codigo
TipoMantenimiento_Descripcion

Stg_TipoLavado
TipoLavado_SKey
TipoLavado_Codigo
TipoLavado_Descripcion

Stg_Lavado
Lavado_SKey
Vehiculo_Placa
Vehiculo_SKey
TipoLavado_Codigo
TipoLavado_SKey
Proveedor_Codigo
Proveedor_SKey
Tiempo_SKey
Lavado_Fecha
Lavado_Costo

Stg_Tiempo
Tiempo_SKey
Tiempo_Fecha
Tiempo_Anio
Tiempo_Mes
Tiempo_Bimestre
Tiempo_Trimestre
Tiempo_Semestre
Tiempo_DescripcionMes
Tiempo_DescripcionBimestre
Tiempo_DescripcionTrimestre
Tiempo_DescripcionSemestre

Stg_Rendimiento
Rendimiento_SKey
Vehiculo_Placa
Vehiculo_SKey
Tiempo_SKey
Rendimiento_Fecha
Rendimiento_GalonConsu...
Rendimiento_KilometroRec...
Rendimiento_Rendimiento...

Stg_TipoParchadoLlanta
TipoParchadoLlanta_SKey
TipoParchadoLlanta_Codigo
TipoParchadoLlanta_Descripcion

Stg_TipoCombustible
TipoCombustible_SKey
TipoCombustible_Codigo
TipoCombustible_Descripcion

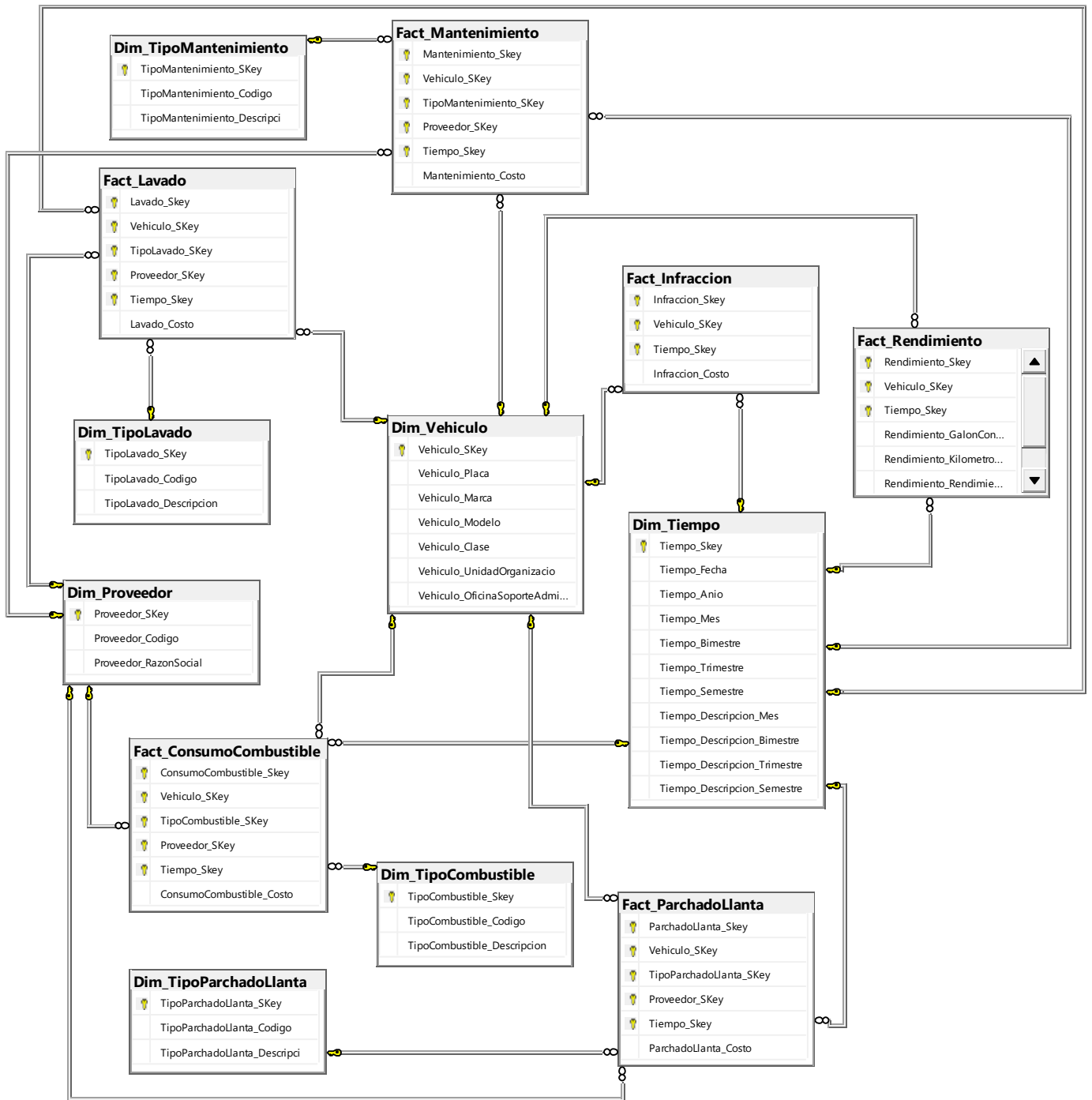
Stg_ParchadoLlanta
ParchadoLlanta_SKey
Vehiculo_Placa
Vehiculo_SKey
TipoParchadoLlanta_Codigo
TipoParchadoLlanta_SKey
Proveedor_Codigo
Proveedor_SKey
Tiempo_SKey
ParchadoLlanta_Fecha
ParchadoLlanta_Costo

Stg_Mantenimiento
Mantenimiento_SKey
Vehiculo_Placa
Vehiculo_SKey
TipoMantenimiento_Codigo
TipoMantenimiento_SKey
Proveedor_Codigo
Proveedor_SKey
Tiempo_SKey
Mantenimiento_Fecha
Mantenimiento_Costo

Stg_Proveedor
Proveedor_SKey
Proveedor_Codigo
Proveedor_RazonSocial

Stg_Infraccion
Infraccion_SKey
Vehiculo_Placa
Vehiculo_SKey
Tiempo_SKey
Infraccion_Fecha
Infraccion_Costo

Base de Datos – Data Mart



Proceso de Transformación

Tablas de Dimensiones

Vehículo

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
				Generado	Dim_Vehiculo	Vehiculo_SKey	int identity
Stg_Vehiculo	Vehiculo_Placa	Alfanumérico (10)	Placa del Vehículo	Copia	Dim_Vehiculo	Vehiculo_Placa	varchar(10)
Stg_Vehiculo	Vehiculo_Marca	Alfanumérico (25)	Marca del Vehículo	Copia	Dim_Vehiculo	Vehiculo_Marca	varchar(25)
Stg_Vehiculo	Vehiculo_Modelo	Alfanumérico (40)	Modelo del Vehículo	Copia	Dim_Vehiculo	Vehiculo_Modelo	varchar(40)
Stg_Vehiculo	Vehiculo_Clase	Alfanumérico (30)	Clase del Vehículo	Copia	Dim_Vehiculo	Vehiculo_Clase	varchar(30)
Stg_Vehiculo	Vehiculo_UnidadOrganizacional	Alfanumérico (80)	Unidad Organizacional a la que pertenece el Vehículo	Copia	Dim_Vehiculo	Vehiculo_Unidad Organizacional	varchar(80)
Stg_Vehiculo	Vehiculo_OficinaSoporteAdministrativa	Alfanumérico (80)	Oficina de Soporte Administrativo a	Copia	Dim_Vehiculo	Vehiculo_Oficina SoporteAdministrativa	varchar(80)

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
			la que pertenece el Vehículo				

Proveedor

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
				Generado	Dim_Proveedor	Proveedor_SKey	int identity
Stg_Proveedor	Proveedor_Codigo	Alfanumérico (5)	Código del Proveedor	Copia	Dim_Proveedor	Proveedor_Codigo	varchar(5)
Stg_Proveedor	Proveedor_RazonSocial	Alfanumérico (50)	Razón social del proveedor	Copia	Dim_Proveedor	Proveedor_RazonSocial	varchar(50)

Tipo de Combustible

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
				Generado	Dim_TipoCombustible	TipoCombustible_SKey	int identity

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_TipoCombustible	TipoCombustible_Codigo	Alfanumérico (5)	Código del Tipo de Combustible	Copia	Dim_TipoCombustible	TipoCombustible_Codigo	varchar(5)
Stg_TipoCombustible	TipoCombustible_Descripcion	Alfanumérico (30)	Descripción del Tipo de Combustible	Copia	Dim_TipoCombustible	TipoCombustible_Descripcion	varchar(30)

Tipo de Lavado

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
				Generado	Dim_TipoLavado	TipoLavado_SKey	int identity
Stg_TipoLavado	TipoLavado_Codigo	Alfanumérico (5)	Código del Tipo de Lavado	Copia	Dim_TipoLavado	TipoLavado_Codigo	varchar(5)
Stg_TipoLavado	TipoLavado_Descripcion	Alfanumérico (30)	Descripción del Tipo de Lavado	Copia	Dim_TipoLavado	TipoLavado_Descripcion	varchar(30)

Tipo de Mantenimiento

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
				Generado	Dim_TipoMantenimiento	TipoMantenimiento_Skey	int identity
Stg_TipoMantenimiento	TipoMantenimiento_Codigo	Alfanumérico (5)	Código del Tipo de Mantenimiento	Copia	Dim_TipoMantenimiento	TipoMantenimiento_Codigo	varchar(5)
Stg_TipoMantenimiento	TipoMantenimiento_Descripcion	Alfanumérico (30)	Descripción del Tipo de Mantenimiento	Copia	Dim_TipoMantenimiento	TipoMantenimiento_Descripcion	varchar(30)

Tipo de Parchado de Llanta

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
				Generado	Dim_TipoParchadoLlanta	TipoParchadoLlanta_Skey	int identity
Stg_TipoParchadoLlanta	TipoParchadoLlanta_Codigo	Alfanumérico (5)	Código del Tipo de ParchadoLlanta	Copia	Dim_TipoParchadoLlanta	TipoParchadoLlanta_Codigo	varchar(5)

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_TipoParchadoLlanta	TipoParchadoLlanta_Descripcion	Alfanumérico (30)	Descripción del Tipo de ParchadoLlanta	Copia	Dim_TipoParchadoLlanta	TipoParchadoLlanta_Descripcion	varchar(30)

Tablas de Hechos

Consumo de Combustible

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_ConsumoCombustible	ConsumoCombustible_SKey	Numérico (5,0)	Clave de Consumo de Combustible	Copia	Fact_ConsumoCombustible	ConsumoCombustible_SKey	Numérico (5,0)
Stg_ConsumoCombustible	Vehiculo_SKey	Numérico (5,0)	Clave de Vehículo	Copia	Fact_ConsumoCombustible	Vehiculo_SKey	Numérico (5,0)
Stg_ConsumoCombustible	TipoCombustible_SKey	Numérico (5,0)	Clave de Tipo de Combustible	Copia	Fact_ConsumoCombustible	TipoCombustible_SKey	Numérico (5,0)

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_ConsumoCombustible	Proveedor_SKey	Numérico (5,0)	Clave de Proveedor	Copia	Fact_ConsumoCombustible	Proveedor_SKey	Numérico (5,0)
Stg_ConsumoCombustible	Tiempo_SKey	Numérico (5,0)	Clave de Tiempo	Copia	Fact_ConsumoCombustible	Tiempo_SKey	Numérico (5,0)
Stg_ConsumoCombustible	ConsumoCombustible_Costo	Numérico (15,2)	Costo por consumo de combustible	Copia	Fact_ConsumoCombustible	ConsumoCombustible_Costo	Numérico (15,2)

Infracción

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_Infraccion	Infraccion_SKey	Numérico (5,0)	Clave de Infracción	Copia	Fact_Infraccion	Infraccion_SKey	Numérico (5,0)
Stg_Infraccion	Vehiculo_SKey	Numérico (5,0)	Clave de Vehículo	Copia	Fact_Infraccion	Vehiculo_SKey	Numérico (5,0)
Stg_Infraccion	Tiempo_SKey	Numérico (5,0)	Clave de Tiempo	Copia	Fact_Infraccion	Tiempo_SKey	Numérico (5,0)

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_Infraccion	Infraccion_Costo	Numérico (15,2)	Costo por infracción	Copia	Fact_Infraccion	Infraccion_Costo	Numérico (15,2)

Lavado

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_Lavado	Lavado_SKey	Numérico (5,0)	Clave de Lavado	Copia	Fact_Lavado	Lavado_SKey	Numérico (5,0)
Stg_Lavado	Vehiculo_SKey	Numérico (5,0)	Clave de Vehículo	Copia	Fact_Lavado	Vehiculo_SKey	Numérico (5,0)
Stg_Lavado	TipoLavado_SKey	Numérico (5,0)	Clave de Tipo de Lavado	Copia	Fact_Lavado	TipoLavado_SKey	Numérico (5,0)
Stg_Lavado	Proveedor_SKey	Numérico (5,0)	Clave de Proveedor	Copia	Fact_Lavado	Proveedor_SKey	Numérico (5,0)
Stg_Lavado	Tiempo_SKey	Numérico (5,0)	Clave de Tiempo	Copia	Fact_Lavado	Tiempo_SKey	Numérico (5,0)
Stg_Lavado	Lavado_Costo	Numérico (15,2)	Costo por lavado	Copia	Fact_Lavado	Lavado_Costo	Numérico (15,2)

Mantenimiento

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_Mantenimiento	Mantenimiento_SKey	Numérico (5,0)	Clave de Mantenimiento	Copia	Fact_Mantenimiento	Mantenimiento_SKey	Numérico (5,0)
Stg_Mantenimiento	Vehiculo_SKey	Numérico (5,0)	Clave de Vehículo	Copia	Fact_Mantenimiento	Vehiculo_SKey	Numérico (5,0)
Stg_Mantenimiento	TipoMantenimiento_SKey	Numérico (5,0)	Clave de Tipo de Mantenimiento	Copia	Fact_Mantenimiento	TipoMantenimiento_SKey	Numérico (5,0)
Stg_Mantenimiento	Proveedor_SKey	Numérico (5,0)	Clave de Proveedor	Copia	Fact_Mantenimiento	Proveedor_SKey	Numérico (5,0)
Stg_Mantenimiento	Tiempo_SKey	Numérico (5,0)	Clave de Tiempo	Copia	Fact_Mantenimiento	Tiempo_SKey	Numérico (5,0)
Stg_Mantenimiento	Mantenimiento_Costo	Numérico (15,2)	Costo por Mantenimiento	Copia	Fact_Mantenimiento	Mantenimiento_Costo	Numérico (15,2)

Parchado de Llantas

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_ParchadoLlanta	ParchadoLlanta_SKey	Numérico (5,0)	Clave de parchado de llantas	Copia	Fact_ParchadoLlanta	ParchadoLlanta_SKey	Numérico (5,0)
Stg_ParchadoLlanta	Vehiculo_SKey	Numérico (5,0)	Clave de Vehículo	Copia	Fact_ParchadoLlanta	Vehiculo_SKey	Numérico (5,0)
Stg_ParchadoLlanta	TipoParchadoLlanta_SKey	Numérico (5,0)	Clave de Tipo de parchado de llantas	Copia	Fact_ParchadoLlanta	TipoParchadoLlanta_SKey	Numérico (5,0)
Stg_ParchadoLlanta	Proveedor_SKey	Numérico (5,0)	Clave de Proveedor	Copia	Fact_ParchadoLlanta	Proveedor_SKey	Numérico (5,0)
Stg_ParchadoLlanta	Tiempo_SKey	Numérico (5,0)	Clave de Tiempo	Copia	Fact_ParchadoLlanta	Tiempo_SKey	Numérico (5,0)
Stg_ParchadoLlanta	ParchadoLlanta_Costo	Numérico (15,2)	Costo por parchado de llantas	Copia	Fact_ParchadoLlanta	ParchadoLlanta_Costo	Numérico (15,2)

Rendimiento

Tabla Origen	Col. Origen	Tipo de Dato	Descripción de la Columna Origen	Transformación	Tabla Destino	Col. Destino	Tipo de Dato
Stg_Rendimiento	Rendimiento_SKey	Numérico (5,0)	Clave de rendimiento	Copia	Fact_Rendimiento	Rendimiento_SKey	Numérico (5,0)
Stg_Rendimiento	Vehiculo_SKey	Numérico (5,0)	Clave de Vehículo	Copia	Fact_Rendimiento	Vehiculo_SKey	Numérico (5,0)
Stg_Rendimiento	Tiempo_SKey	Numérico (5,0)	Clave de Tiempo	Copia	Fact_Rendimiento	Tiempo_SKey	Numérico (5,0)
Stg_Rendimiento	Rendimiento_GalonConsumido	Numérico (15,2)	Rendimiento vehicular	Copia	Fact_Rendimiento	Rendimiento_GalonConsumido	Numérico (15,2)
Stg_Rendimiento	Rendimiento_KilometroRecorrido	Numérico (15,2)	Rendimiento vehicular	Copia	Fact_Rendimiento	Rendimiento_KilometroRecorrido	Numérico (15,2)
Stg_Rendimiento	Rendimiento_RendimientoVehicular	Numérico (15,2)	Rendimiento vehicular	Calculado = Rendimiento_KilometroRecorrido / Rendimiento_GalonConsumido	Fact_Rendimiento	Rendimiento_RendimientoVehicular	Numérico (15,2)

Herramientas de diseño del Cubo

Matriz Bus

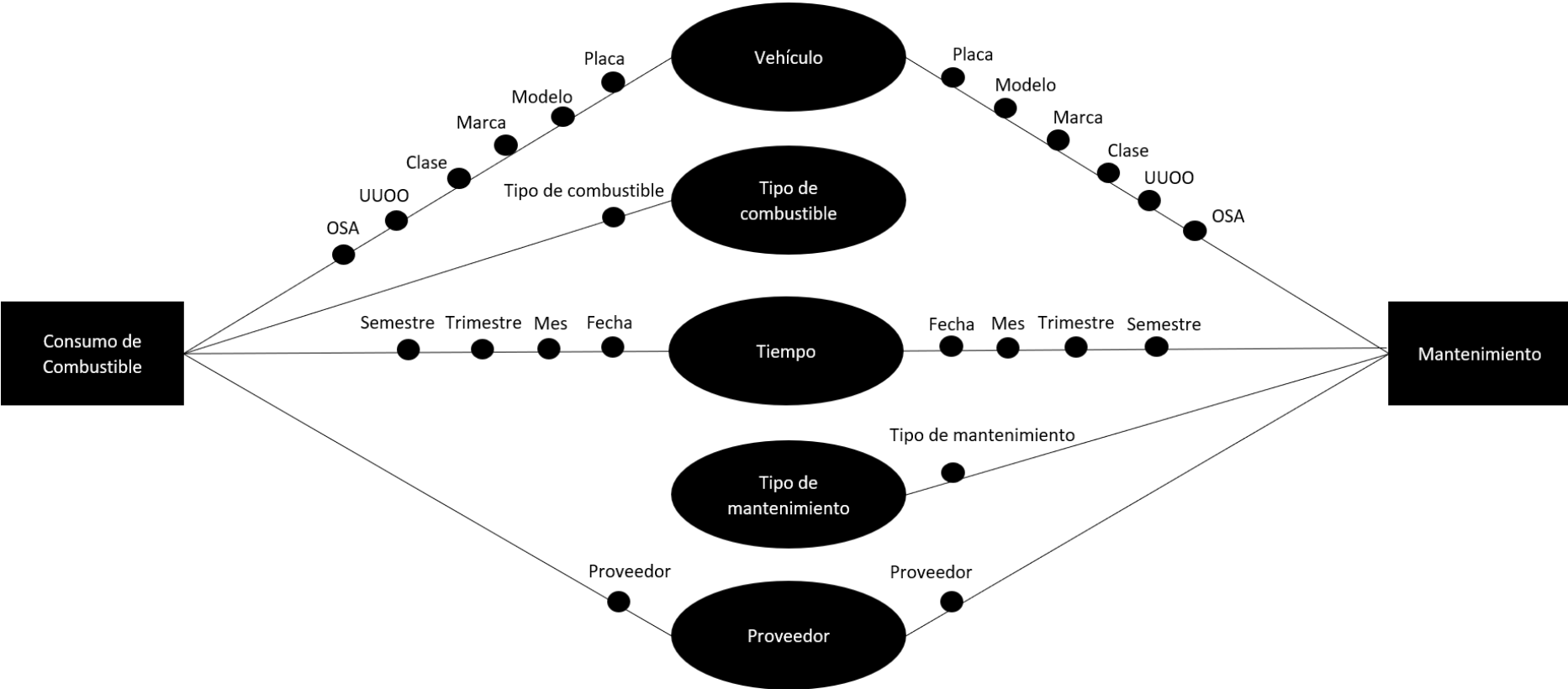
Procesos	Dimensión	Vehículo	Tipo de Combustible	Proveedor	Tipo de Mantenimiento	Tiempo	Tipo de Lavado	Tipo de Parchado de Llantas
	Métrica							
Consumo de Combustible	Costo por consumo de combustible	X	X	X		X		
	Porcentaje del costo por consumo de combustible (calculado)	X	X	X		X		
Mantenimiento	Costo por mantenimiento	X		X	X	X		
	Porcentaje del costo por mantenimiento (calculado)	X		X	X	X		
Infracción	Costo por infracciones	X				X		
	Porcentaje de costo por accidentes o infracciones (calculado)	X				X		
Parchado de Llantas	Costo por llantas	X		X		X		X
	Porcentaje de costo por llantas (calculado)	X		X		X		X

Lavado	Costo por lavado	X				X	X	
	Porcentaje de costo por lavado (calculado)	X				X	X	
Rendimiento	Kilometraje recorrido	X				X		
	Cantidad de galones consumidos	X				X		
	Rendimiento Vehicular	X				X		
	Promedio de rendimiento (calculado)	X				X		

Star-Net

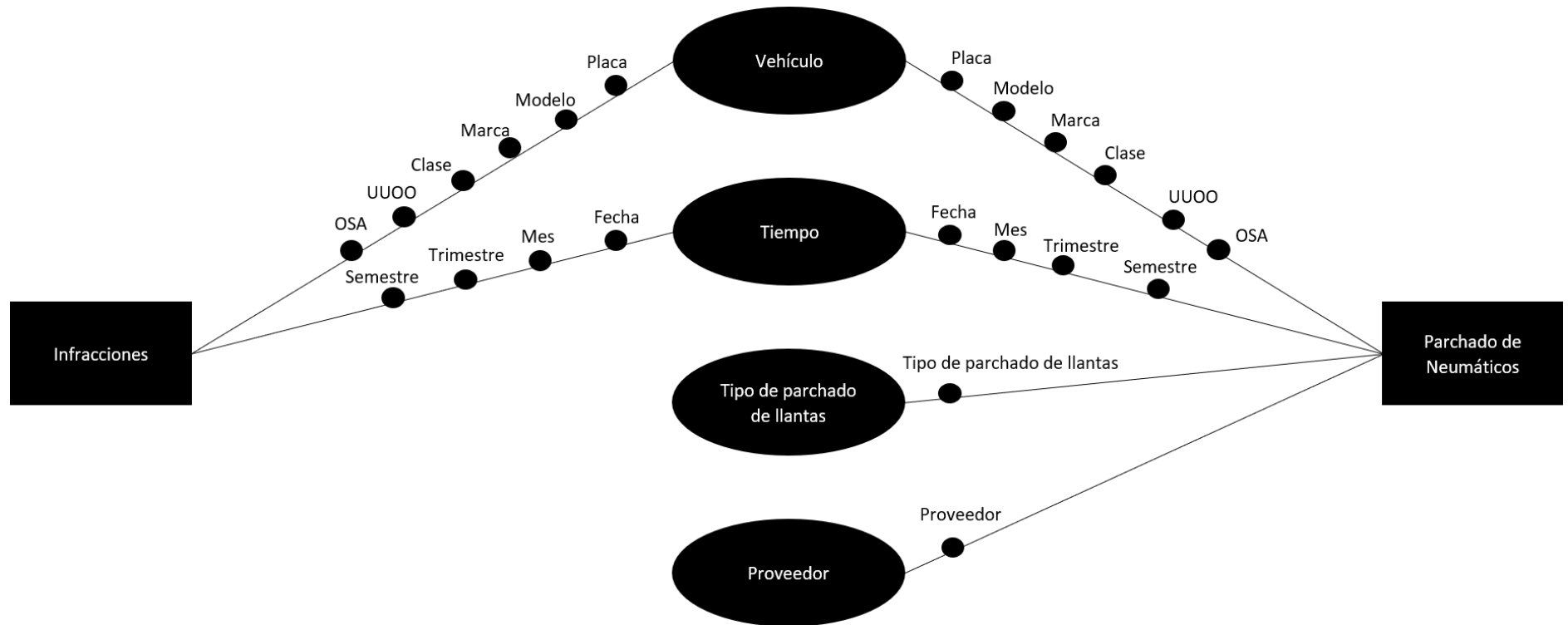
Procesos:

- Consumo de Combustible.
- Mantenimiento.



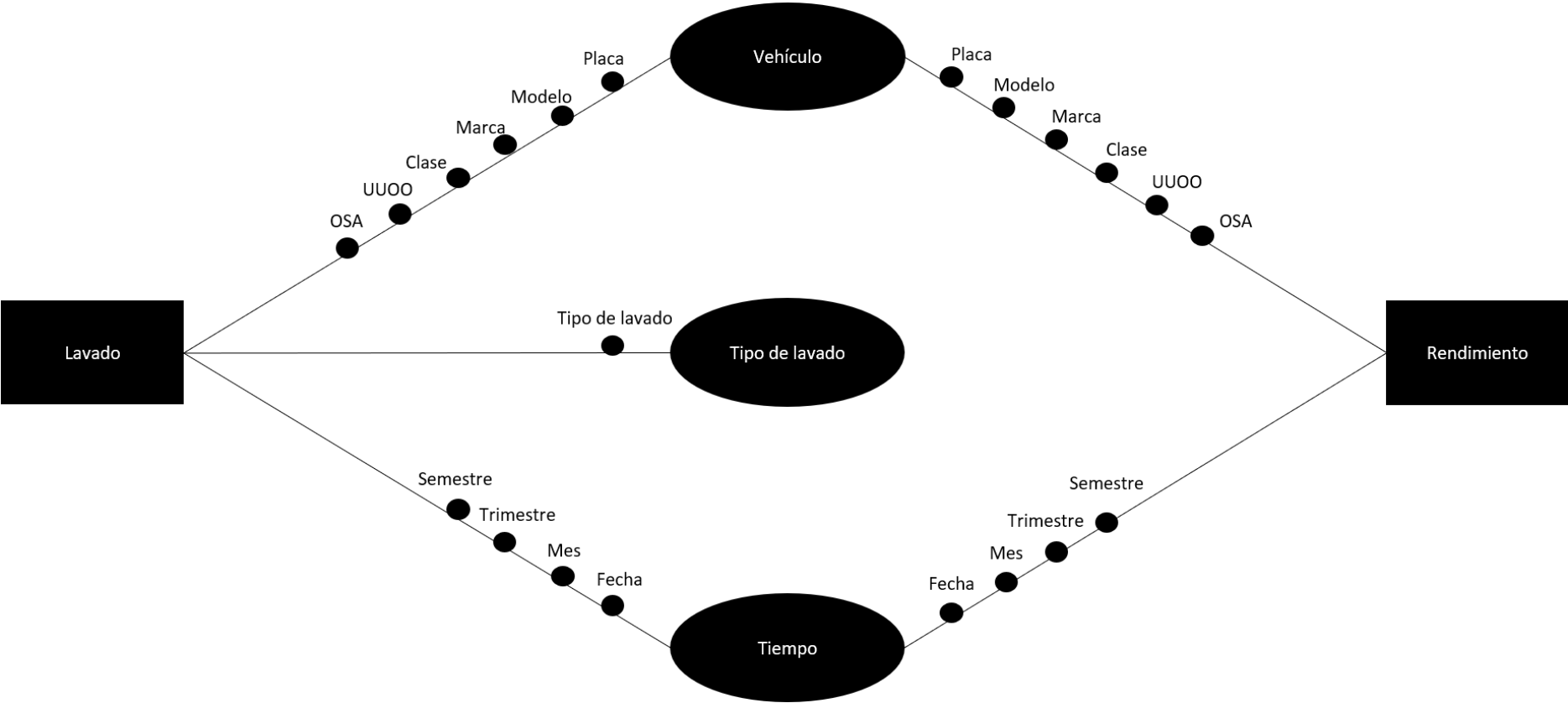
Procesos:

- Infracciones.
- Parchado de LLantas.



Procesos:

- Lavado.
- Rendimiento.



Jerarquías

Dim. Vehículo

J1: Clase_Vehículo

**Oficina de Soporte Administrativo

***Unidad Organizacional

****Clase

*****Placa

Dim. Tiempo

*Semestre

**Trimestre

***Mes

****Fecha

Dim. Tipo de Combustible

*Tipo de Combustible

Dim. Tipo de Mantenimiento

*Tipo de Mantenimiento

Dim. Vehículo

J2: Marca_Vehículo

**Oficina de Soporte Administrativo

***Unidad Organizacional

****Marca

* ****Modelo

*****Placa

Dim. Proveedor

*Proveedor

Dim. Tipo de Lavado

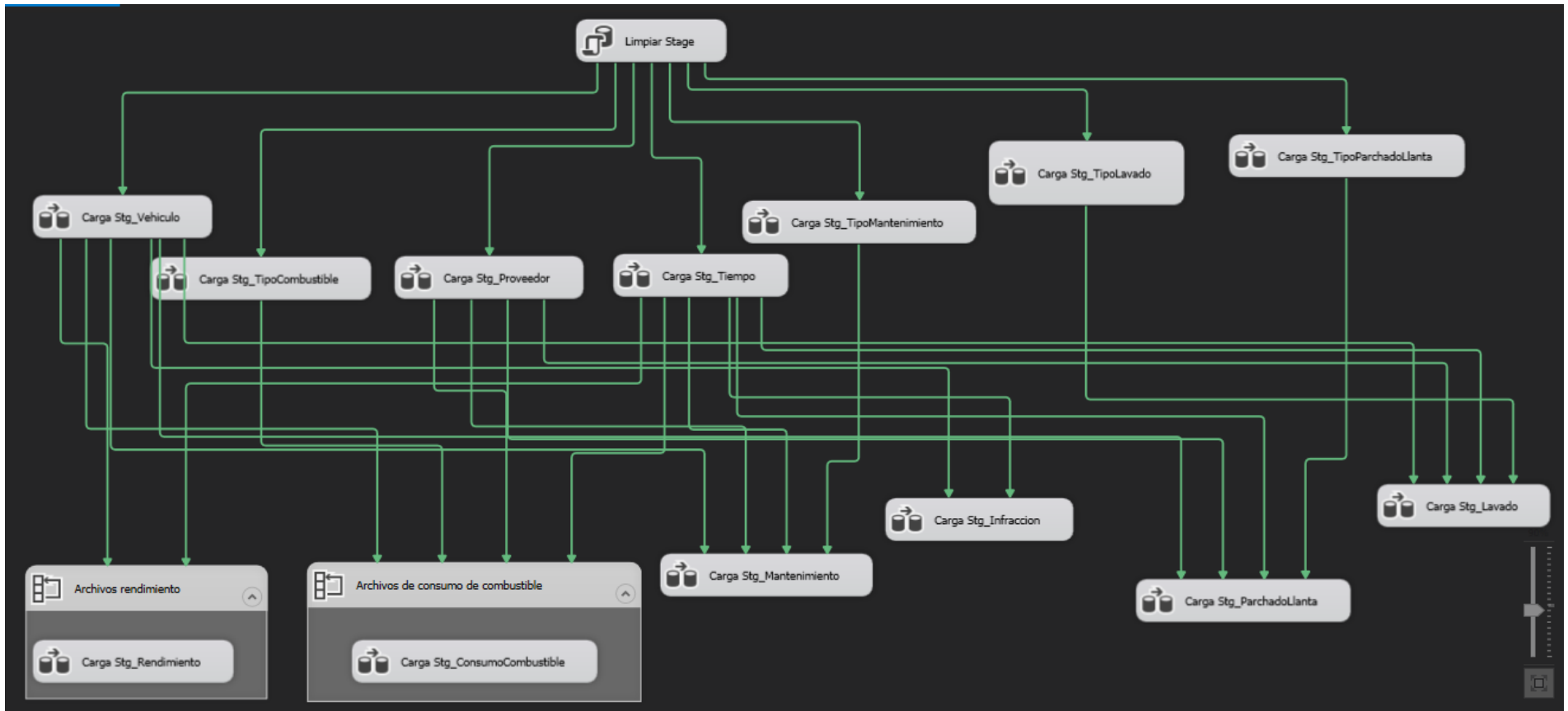
*Tipo de Lavado

Dim. Tipo de Parchado de Llantas

*Tipo de Parchado de Llantas

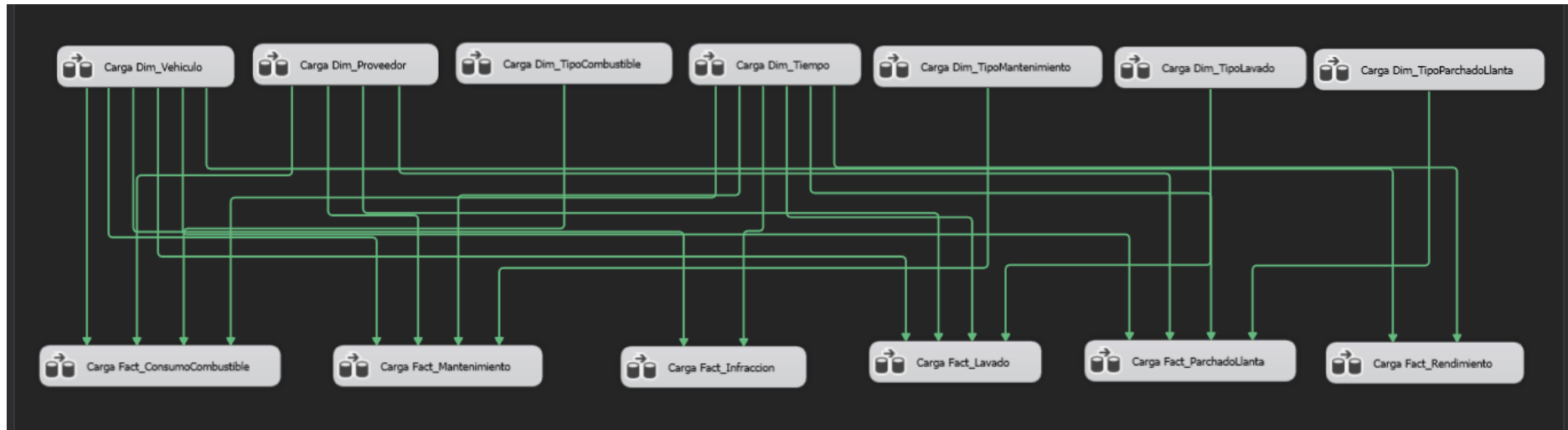
Jerarquías con respecto a las dimensiones, por las cuales se realizarán los análisis dentro del cubo OLAP,

Extracción de fuentes de datos y Transformación en el Data Stage



Proceso de Extracción y Transformación de datos de las fuentes de datos de la gestión de la flota vehicular; se extraen los datos de las fuentes de datos, para proceder con la respectiva limpieza y transformación de los datos, almacenando estos datos procesados en la base de datos intermedia o Data Stage.

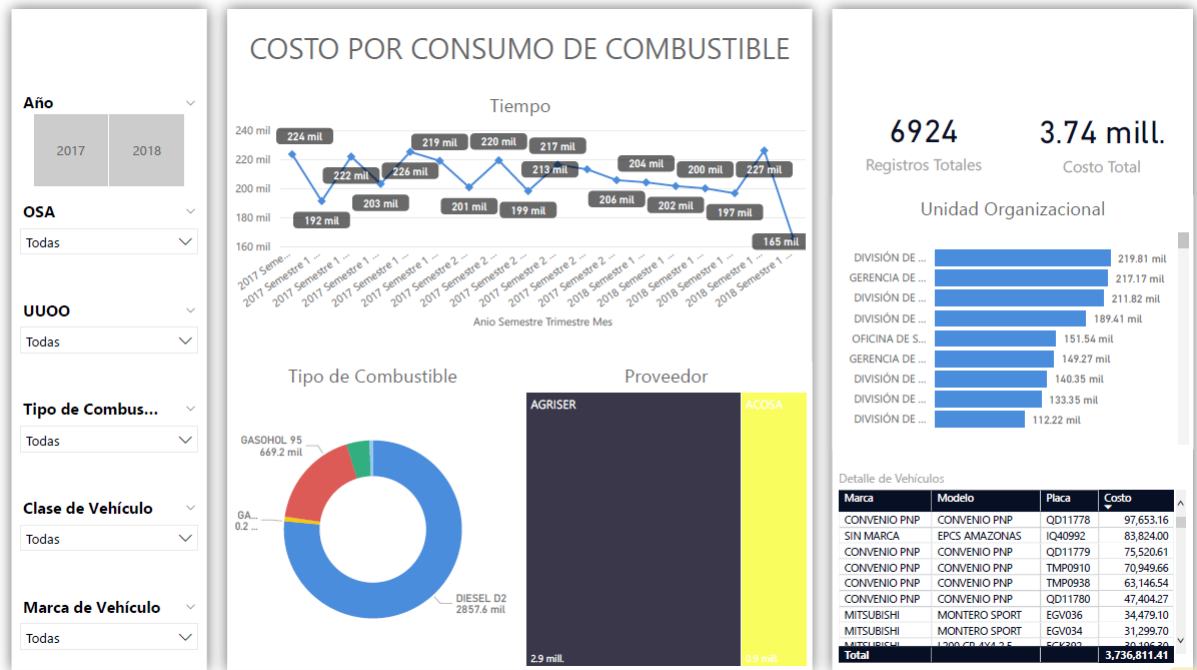
Poblamiento del Data Mart



Proceso de carga de datos al Data Mart de la gestión de la flota vehicular; se extraen los datos del Data Stage en el cual se generó la limpieza y transformación de los datos, para proceder con la respectiva carga, culminando así el proceso de Extracción, Transformación y Carga de los datos.

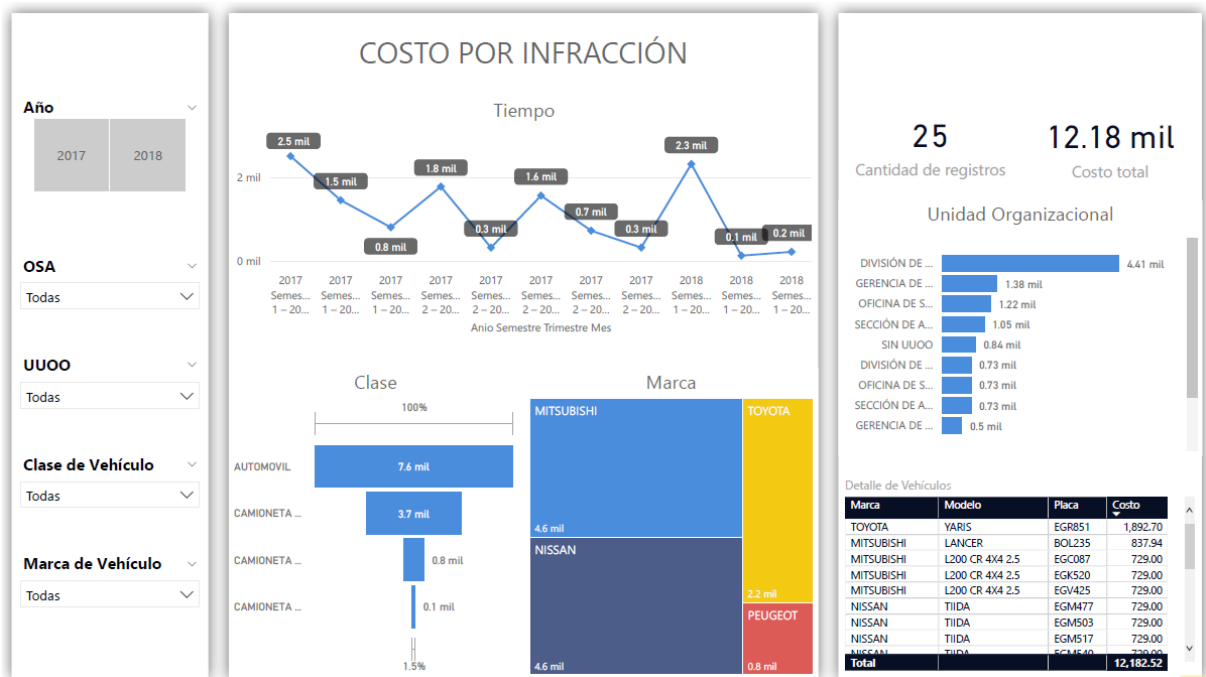
Dashboards

Costo por Consumo de Combustible



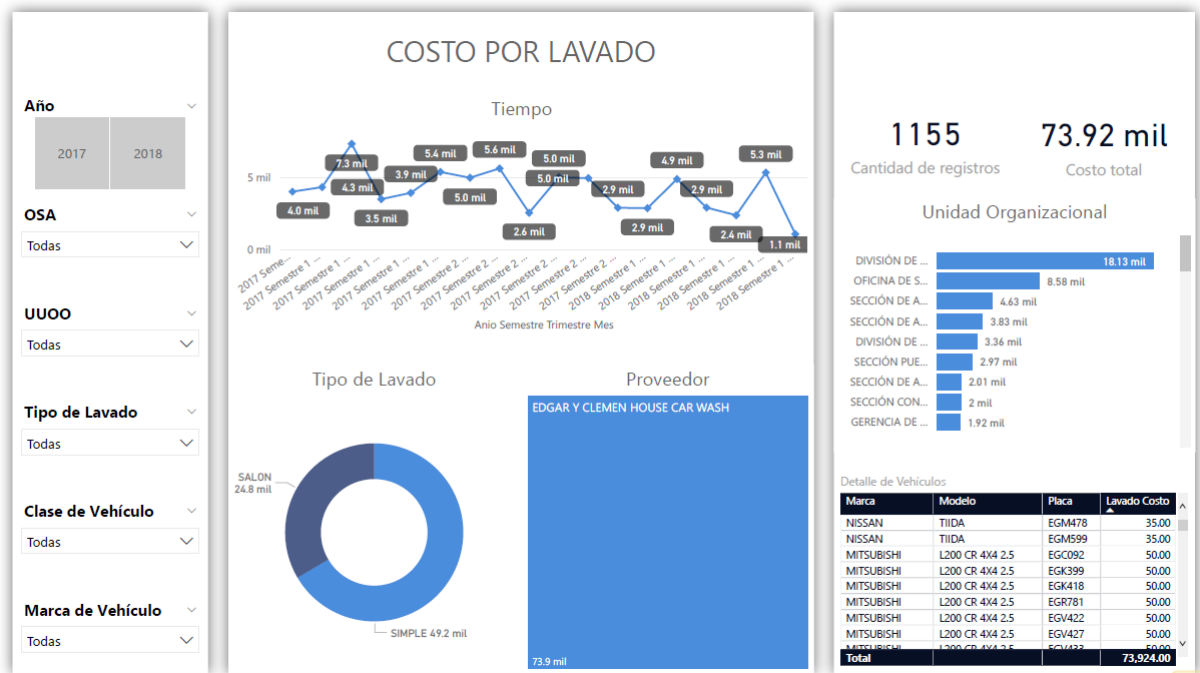
Reporte del costo por combustible analizado por las dimensiones de Vehículo el cual está representado de manera tabular, dimensión Tipo de Combustible el cual esta representado en un grafico de anillos, dimensión Tiempo el cual se puede visualizar en una linea temporal y dimensión Proveedor el cual se puede visualizar en un grafico de proporciones, así mismo se puede visualizar por una de las características de la dimensión Vehiculo la cual es la Unidad Organizacional que esta representada en un grafico de barras horizontal, se tiene un resumen de la cantidad de registros totales existentes y del costo total sumado. Por último se pueden realizar filtros por el Año, OSA(Oficina de Soporte Admnistrativa), UUOO (Unidad Organizacional), Tipo de Combustible, Clase de Vehículo y Marca de Vehículo; estos filtros se verán aplicados en todo el reporte y/o gráficos de Power BI.

Costo por Infracción



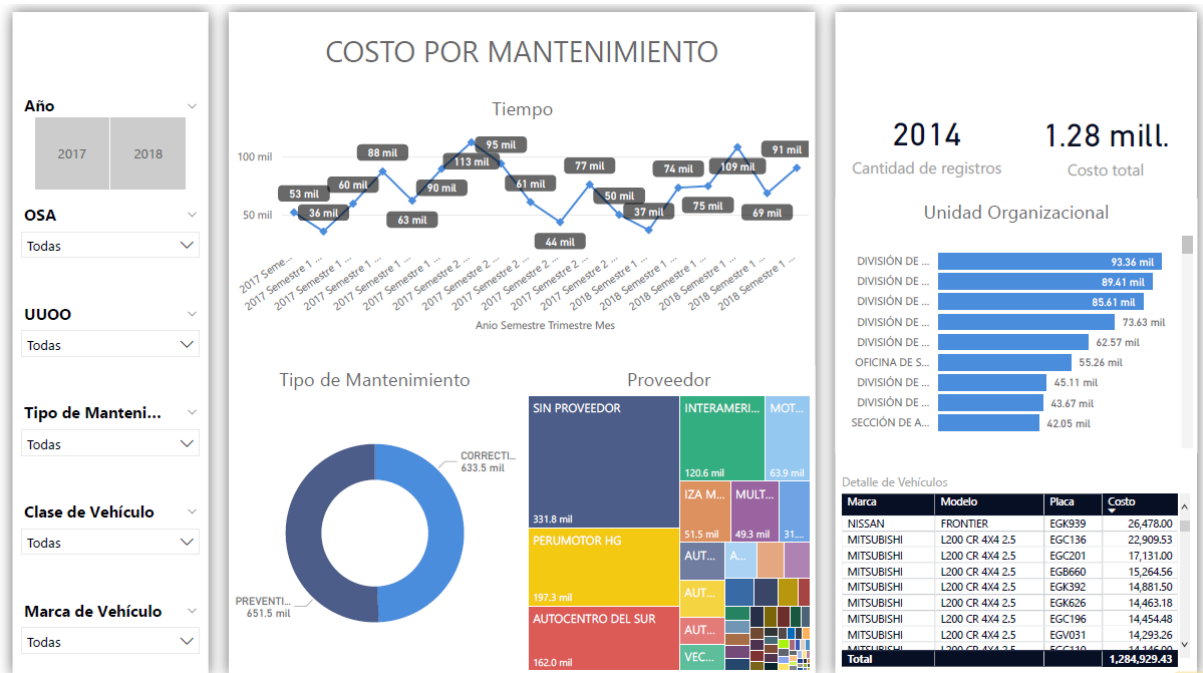
Reporte del costo por infracción analizado por las dimensiones de Vehículo el cual está representado de manera tabular y dimensión Tiempo el cual se puede visualizar en una línea temporal, así mismo se puede visualizar por una de las características de la dimensión Vehículo la cual es la Unidad Organizacional que esta representada en un gráfico de barras horizontal, una de las características de la dimensión Vehículo la cual es la Marca que esta representada en un gráfico de proporciones, y una de las características de la dimensión Vehículo la cual es la Clase que esta representada en un gráfico de barras proporcional, se tiene un resumen de la cantidad de registros totales existentes y del costo total sumado. Por último se pueden realizar filtros por el Año, OSA(Oficina de Soporte Administrativa), UUOO (Unidad Organizacional), Clase de Vehículo y Marca de Vehículo; estos filtros se verán aplicados en todo el reporte y/o gráficos de Power BI.

Costo por Lavado



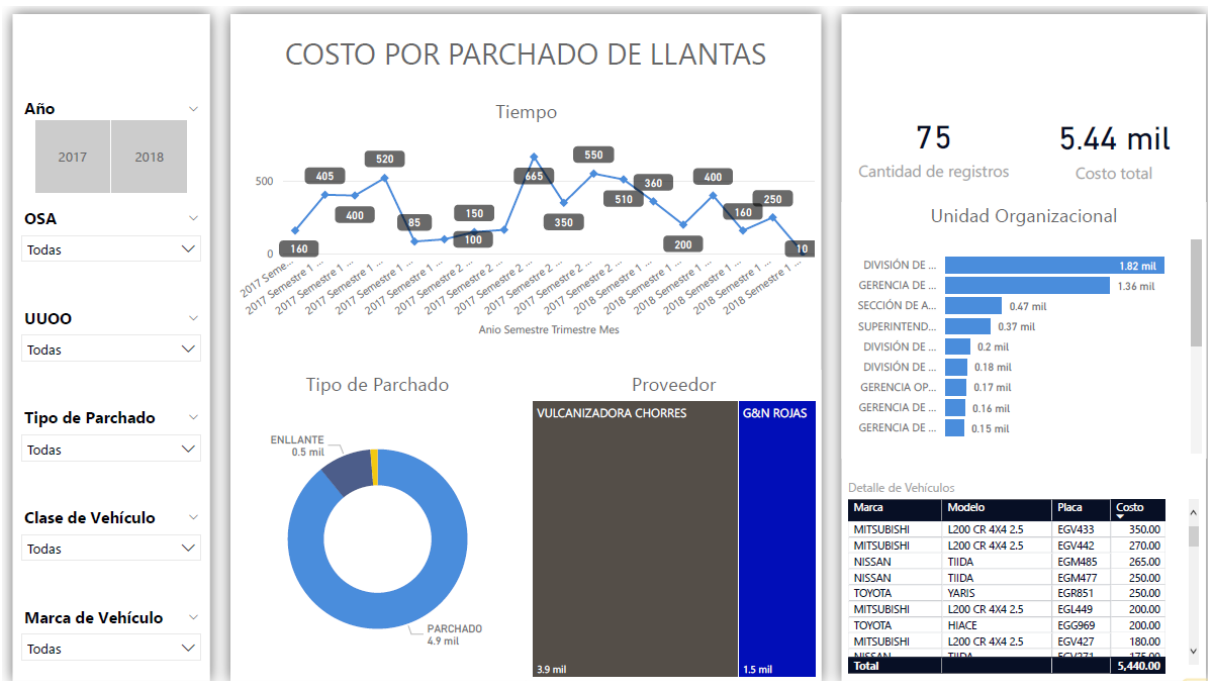
Reporte del costo por lavado analizado por las dimensiones de Vehículo el cual está representado de manera tabular, dimensión Tipo de Lavado el cual está representado en un gráfico de anillos, dimensión Tiempo el cual se puede visualizar en una línea temporal y dimensión Proveedor el cual se puede visualizar en un gráfico de proporciones, así mismo se puede visualizar por una de las características de la dimensión Vehículo la cual es la Unidad Organizacional que está representada en un gráfico de barras horizontal, se tiene un resumen de la cantidad de registros totales existentes y del costo total sumado. Por último se pueden realizar filtros por el Año, OSA (Oficina de Soporte Administrativa), UUOO (Unidad Organizacional), Tipo de Lavado, Clase de Vehículo y Marca de Vehículo; estos filtros se verán aplicados en todo el reporte y/o gráficos de Power BI.

Costo por Mantenimiento



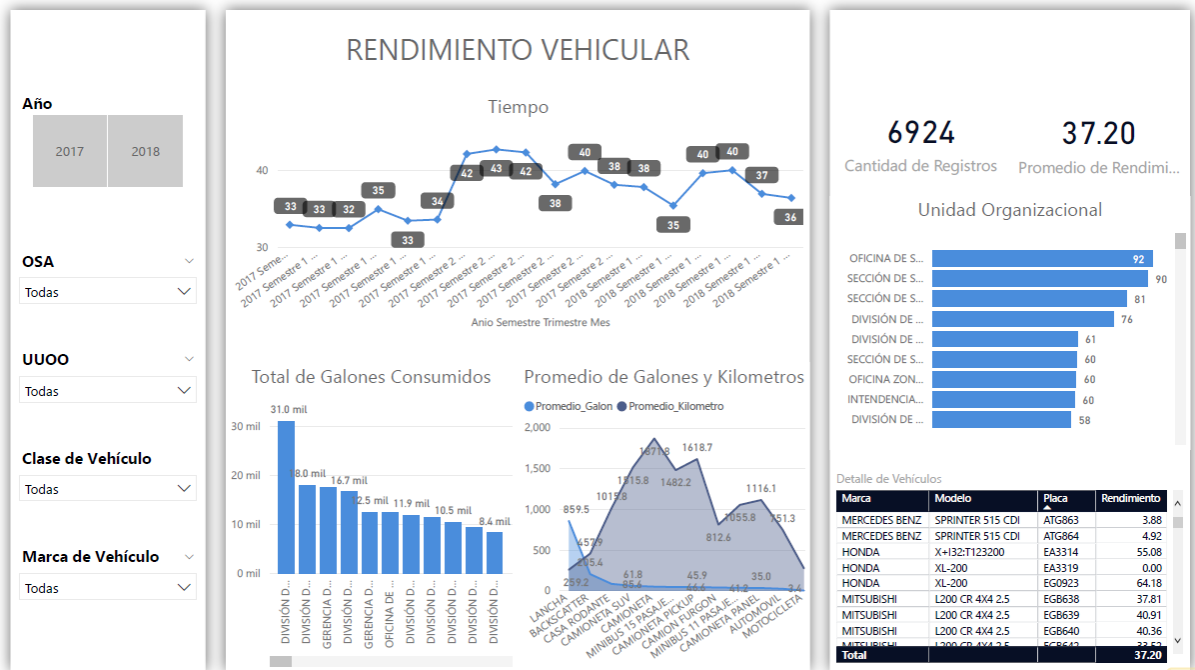
Reporte del costo por lavado analizado por las dimensiones de Vehículo el cual está representado de manera tabular, dimensión Tipo de Mantenimiento el cual esta representado en un grafico de anillos, dimensión Tiempo el cual se puede visualizar en una linea temporal y dimensión Proveedor el cual se puede visualizar en un grafico de proporciones, así mismo se puede visualizar por una de las características de la dimensión Vehículo la cual es la Unidad Organizacional que esta representada en un grafico de barras horizontal, se tiene un resumen de la cantidad de registros totales existentes y del costo total sumado. Por último se pueden realizar filtros por el Año, OSA(Oficina de Soporte Admnistrativa), UUOO (Unidad Organizacional), Tipo de Mantenimiento, Clase de Vehículo y Marca de Vehículo; estos filtros se verán aplicados en todo el reporte y/o gráficos de Power BI.

Costo por Parchado de Llantas



Reporte del costo por lavado analizado por las dimensiones de Vehículo el cual está representado de manera tabular, dimensión Tipo de Parchado de Llantas el cual esta representado en un grafico de anillos, dimensión Tiempo el cual se puede visualizar en una linea temporal y dimensión Proveedor el cual se puede visualizar en un grafico de proporciones, así mismo se puede visualizar por una de las características de la dimensión Vehículo la cual es la Unidad Organizacional que esta representada en un grafico de barras horizontal, se tiene un resumen de la cantidad de registros totales existentes y del costo total sumado. Por último se pueden realizar filtros por el Año, OSA(Oficina de Soporte Admnistrativa), UUOO (Unidad Organizacional), Tipo de Parchado de Llantas, Clase de Vehículo y Marca de Vehículo; estos filtros se verán aplicados en todo el reporte y/o gráficos de Power BI.

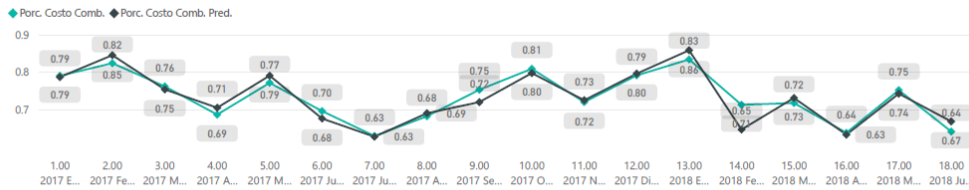
Rendimiento Vehicular



Reporte del costo por infracción analizado por las dimensiones de Vehículo el cual está representado de manera tabular y dimensión Tiempo el cual se puede visualizar en una línea temporal, así mismo se puede visualizar por una de las características de la dimensión Vehículo la cual es la Unidad Organizacional que esta representada en un gráfico de barras horizontal, una de las medidas de Rendimiento Vehicular la cual es la Galones Consumidos que esta representada en un gráfico de de barras verticalarl, y una de las características de las medidas de Rendimiento Vehicular la cual es la Kilometros Recorridos que esta representada en un gráfico de áreas, se tiene un resumen de la cantidad de registros totales existentes y del rendimiento promediado. Por último se pueden realizar filtros por el Año, OSA(Oficina de Soporte Admnistrativa), UUOO (Unidad Organizacional), Clase de Vehículo y Marca de Vehículo; estos filtros se verán aplicados en todo el reporte y/o gráficos de Power BI.

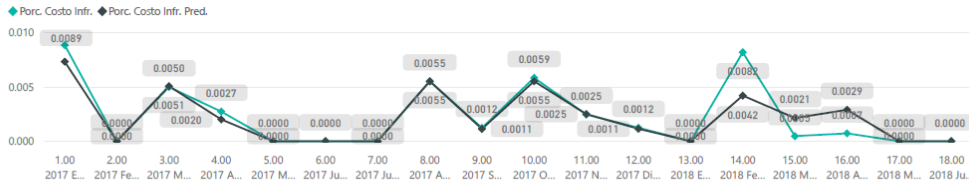
Predicción de Indicadores

PORCENTAJE DE COSTO POR CONSUMO DE COMBUSTIBLE



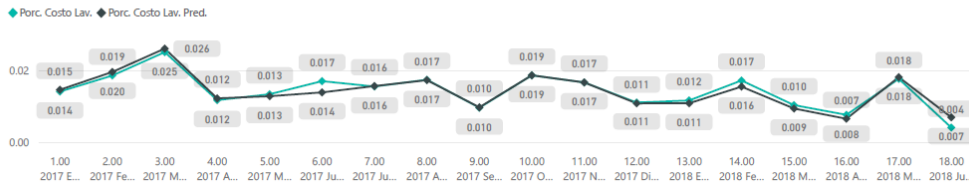
152.12
Costo de Combustible Pred.

PORCENTAJE DE COSTO POR INFRACCIÓN



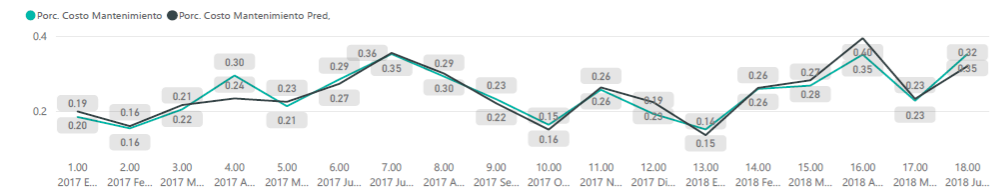
298.09
Costo de Infracción Pred.

PORCENTAJE DE COSTO POR LAVADO



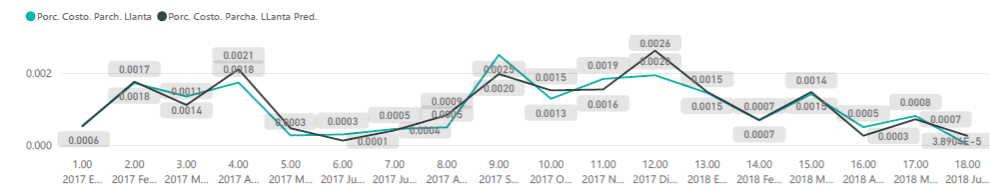
35.00
Costo de Lavado Pred.

PORCENTAJE DE COSTO POR MANTENIMIENTO



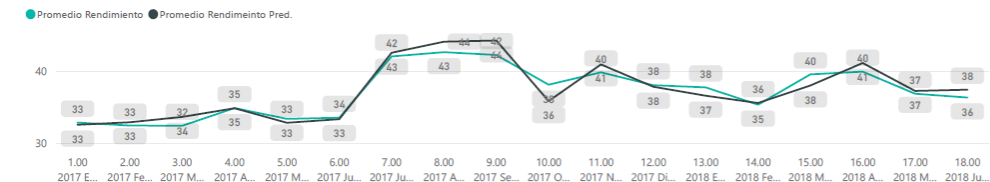
368.22
Costo de Mantenimiento Pred.

PORCENTAJE DE COSTO POR PARCHADO DE LLANTAS



71.33
Costo de Parch. Llanta Pred.

RENDIMIENTO VEHICULAR



57.29
Rendimiento Vehicular Pred.

Predicciones de los costos por combustible, infracciones, lavado, mantenimiento, parchado de llantas y del rendimiento vehicular, emitidos por el modelo de

machine learning en base a los datos de los 18 meses de la operatividad de la flota vehicular

Anexo 4 Autorización de publicación de Tesis



UNIVERSIDAD CÉSAR VALLEJO

**FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

Autorización de Publicación en Repositorio Institucional

Yo, VERGARA PINEDO JUAN LUIS identificado con DNI N° 76295406, (respectivamente) estudiante de la FACULTAD DE INGENIERÍA Y ARQUITECTURA y de la escuela profesional de INGENIERÍA DE SISTEMAS de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA ESTE, autorizo (X), no autorizo () la divulgación y comunicación pública de mi Tesis: "Desarrollo de Inteligencia de Negocios con un modelo de Machine Learning para la gestión de la flota vehicular".

En el Repositorio Institucional de la Universidad César Vallejo, según esta estipulado en el Decreto Legislativo 822, Ley sobre Derecho de Autor, Art. 23 y Art. 33.

Fundamentación en caso de NO autorización:

--

SAN JUAN DE LURIGANCHO, 22 de Diciembre del 2021

Apellidos y Nombres del Autor	Firma
VERGARA PINEDO JUAN LUIS DNI: 76295406 ORCID 0000-0002-9233-4029	Firmado digitalmente por: JVERGARAP7 el 22-12- 2021 09:29:48

Código documento Trilce: INV - 0459735



Anexo 5 Autorización de la versión final del trabajo de investigación



UNIVERSIDAD CÉSAR VALLEJO

AUTORIZACIÓN DE LA VERSIÓN FINAL DEL TRABAJO DE INVESTIGACIÓN

CONSTE POR EL PRESENTE EL VISTO BUENO QUE OTORGA EL ENCARGADO DE INVESTIGACIÓN DE LA ESCUELA DE INGENIERÍA DE SISTEMAS, MG. MARÍA ACUÑA MELÉNDEZ A LA RECEPCIÓN DE LA DOCUMENTACIÓN SOLICITADA PARA LA ENTREGA DE LA VERSIÓN FINAL DEL TRABAJO DE INVESTIGACIÓN QUE PRESENTA:

VERGARA PINEDO JUAN LUIS

INFORME TÍTULADO:

DESARROLLO DE INTELIGENCIA DE NEGOCIOS CON UN MODELO DE MACHINE LEARNING PARA LA GESTIÓN DE LA FLOTA VEHICULAR

PARA OBTENER EL TÍTULO O GRADO DE:

INGENIERO DE SISTEMAS

SUSTENTADO EN FECHA: San Juan de Lurigancho, 16 de Diciembre del 2018

NOTA O MENCIÓN: 13 (Trece)



Mg. María Acuña Meléndez
CP de Ingeniería de Sistemas campus Lima Este