UNIVERSIDAD CENFOTEC

Maestría en Tecnologías en Bases de Datos

Documento final de Proyecto de Investigación Aplicada 2

Estimación del cumplimiento de entrega de productos al cliente final para una empresa privada utilizando aprendizaje automático

Elaborado por:
Brenes Núñez Marilyn

Marzo, 2022

Declaratoria de derechos de autor

Se declara que el presente proyecto de investigación fue realizado por la autora por Marilyn Brenes Núñez, fundamentando los diferentes capítulos del trabajo en diferentes fuentes bibliográficas, literatura citada, las cuales tienen su respectiva referencia, respetando los derechos de autor de dichos trabajos. De igual manera se tomó como referencia datos obtenidos de diferentes metodologías de trabajo, en las cuales se reflejan los comentarios de los profesionales expertos.

Se solicita el congelamiento de publicación y reproducción de este trabajo por un año, una vez sobrepasado este periodo se autoriza la reproducción total o parcial para ser usado como referencia de trabajos futuros de tipo académico y científico, en este caso, se solicita incorporar la referencia de este trabajo respetando los derechos de la autora.

Dedicatoria

A mi madre

Eres mi heroína, tu amor y apoyo incondicional
son la fuerza que me impulsa a seguir adelante.

A mi tía

"Ánimo, ¡usted puede!,
algún día te sorprenderás de lo capaz que eres."

A mi abuelita

"…en un cielo lleno de estrellas, creo verte en una de ellas…"

A las chicas en tecnología

Sean auténticas y crean en ese poder con valentía,
esa es la magia que les señalará la ruta de sus sueños
y si a eso le agregas un salto de fe,
pueda que termine siendo la aventura más increíble de sus vidas.

Agradecimientos

Empresa patrocinadora

Especial agradecimiento por su generosidad, apoyo
y compromiso para el éxito de este trabajo.

Señor Silvio Alvim

Gurú en temas de Supply Chain,
gracias por sus tan valiosos aportes en la investigación.

Profesores José Alberto Cabezas y Luis Carlos Naranjo

Millones de gracias por su gran disposición, sugerencias
y sobre todo por el cariño en el trabajo que realizan.

Universidad Cenfotec
Carrera de Postgrado
Maestría en Tecnología de Bases de Datos

TRIBUNAL EXAMINADOR

Este proyecto fue aprobado por el Tribunal Examinador de la carrera: **Maestría en Tecnología de Bases de Datos,** requisito para optar por el título de grado de **Maestría,** para el estudiante: **Brenes Núñez Marilyn.**

JOSE ALBERTO CABEZAS JAIKEL (FIRMA) Digitally signed by JOSE ALBERTO CABEZAS JAIKEL (FIRMA) Date: 2022.03.09 09:15:37 -06'00'

MBD. José Cabezas Jaikel
Tutor

LUIS CARLOS NARANJO ZELEDON (FIRMA) Firmado digitalmente por LUIS CARLOS NARANJO ZELEDON (FIRMA) Fecha: 2022.03.10 13:42:30 -06'00'

Dr. Luis Naranjo Zeledón
Lector 1

IGNACIO TREJOS ZELAYA (FIRMA) Firmado digitalmente por IGNACIO TREJOS ZELAYA (FIRMA) Fecha: 2022.03.10 18:57:03 -06'00'

M. Sc. Ignacio Trejos Zelaya
Lector 2

San José, Costa Rica, 08 de marzo de 2022

Firmada digitalmente, de conformidad con la Ley de Certificados, Firmas Digitales y Documentos Electrónicos N° 8454, destacando el artículo 9°-

Tabla de Contenidos

Tabla de Contenidos iv
Lista de Figuras viii
Capítulo 1 Introducción 1
1.1 Generalidades 1
1.2 Antecedentes del problema 1
1.3 Definición y descripción del problema 3
1.4 Justificación 4
1.5 Viabilidad 7
1.5.1 Punto de vista técnico. 7
1.5.2 Punto de vista operativo. 8
1.5.3 Punto de vista económico. 8
1.6 Objetivos 9
1.6.1 Objetivo general. 9
1.6.2 Objetivos específicos. 9
1.7 Alcances y Limitaciones 10
1.7.1 Alcances. 10
1.7.2 Limitaciones. 10
1.8 Marco de referencia organizacional y socioeconómico 12
1.9 Estado de la cuestión. 12
1.9.1 Planificación de la revisión. 12
1.9.2 Formulación de la pregunta. 12
1.9.3 Selección de fuentes. 15
1.9.4 Selección de los estudios. 18
1.9.5 Ejecución de la revisión. 20
1.9.6 Resumen de los resultados. 30
Capítulo 2 Marco Conceptual 34
2.1 Cadena de suministro (Supply Chain SC) 35
2.1.1 Riesgo en cadena de suministros (Supply Chain Risk SCR) 36

2.2 Inteligencia artificial (AI).....38
2.2.1 Aprendizaje automático (Machine Learning ML).....38
2.2.2 Regresión lineal.....39
2.2.3 Regresión logística.....39
2.2.4 Bosque aleatorio.....40
2.2.5 Redes neuronales artificiales (ANN).....40
Capítulo 3 Marco Metodológico.....43
3.1 Tipo de investigación.....43
3.2 Alcance investigativo.....43
3.3 Enfoque.....44
3.4 Población y muestreo.....51
3.5 Instrumentos de recolección de datos.....51
3.6 Técnicas para análisis de información.....54
Capítulo 4 Análisis del diagnóstico.....55
4.1 Técnica #1 Lluvia de ideas.....55
4.2 Técnica #4 Mapa conceptual.....57
4.3 Técnica #5 Mapa de flujo de valor.....61
4.4 Técnica #2 SIPOC.....61
Capítulo 5 Propuesta de solución.....63
5.1 Modelo predictivo del cumplimiento de entrega.....63
5.1.1 Entendimiento del negocio.....64
5.1.2 Comprensión de los datos.....64
5.1.3 Preparación de los datos.....67
5.1.4 Exploración de los datos.....76
5.1.5 Modelado.....92
5.1.6 Evaluación.....96
5.2 Propuesta de alteración a la metodología del cálculo de tiempos de entrega.....97
Capítulo 6 Conclusiones y recomendaciones.....100
6.1 Conclusiones.....100

6.1.1 Conclusión objetivo #1. 101
6.1.2 Conclusión objetivo #2. 101
6.1.3 Conclusión objetivo #3. 102
6.1.4 Conclusión objetivo #4. 102
6.1.5 Conclusión objetivo #5. 103
6.1.6 Conclusión objetivo general. 103
6.2 Recomendaciones 104
6.2.1 Recomendaciones técnicas. 104
6.2.2 Recomendaciones operativas. 105
Capítulo 7 Reflexiones Finales 107
Capítulo 8 Trabajos a futuro 109
Referencias bibliográficas 111
Apéndices 115

Lista de Tablas

Tabla 1 Nivel de conocimiento de la investigadora 7
Tabla 2 Escala de salarios máximos de acuerdo con la página web tusalario.org de Costa Rica en colones, 2020. 8
Tabla 3 Recopilación de costos mensuales del proyecto para patrocinador, 2021. 9
Tabla 4 Recopilación de costos mensuales del proyecto para investigador, 2021. 9
Tabla 5 Palabras clave 13
Tabla 6 Pasos de la identificación de fuentes. 16
Tabla 7 Criterio de inclusión y exclusión de estudios 19
Tabla 8 Lista de estudios a evaluar. 21
Tabla 9 Valoración de los estudios publicados por revista 22
Tabla 10 Documentos, conferencias o libros incluidos por juicio experto. 23
Tabla 11 Evaluación de criterios. 24
Tabla 12 Extracción de información, documento J-a. 25
Tabla 13 Extracción de información, documento J-b. 26
Tabla 14 Extracción de información, documento J-c. 27
Tabla 15 Extracción de información, documento J-d. 28
Tabla 16 Parámetros textuales de referencia para gráfico. 31
Tabla 17 Evaluación del acercamiento del tema por pregunta 31
Tabla 18 Rúbrica para determinar el grado de dificultad de la implementación del proyecto. 46
Tabla 19 Rúbrica para determinar el grado de dificultad de la implementación del proyecto. 47
Tabla 20 Definición de las etapas DMAIC. 50
Tabla 21 Descripción de variables 64
Tabla 22 Detalle de las variables usadas por recurso 67
Tabla 23 Comentarios orden A 73
Tabla 24 Limpieza por variable expresado en porcentaje 75
Tabla 25 Técnicas de maquillaje ético 75
Tabla 26 Correlación entre variables cuantitativas. 79
Tabla 27 Variables por modelos 93
Tabla 28 Resultados evaluación de modelo y algoritmo. 96
Tabla 29 Resumen estadístico de factores principales de atraso 97
Tabla 30 Variantes de los tiempos estándar. 97
Tabla 31 Resumen estadístico de variables extra. 98
Tabla 32 Cumplimiento de objetivos utilizando DMAIC 100

Lista de Figuras

Figura 1 Beneficios mensuales de este proyecto. Elaboración propia, 2020. 4
Figura 2 Interacción entre equipos. Elaboración propia, 2020. 5
Figura 3 Distribución de tareas en un día laboral. Elaboración propia, 2021. 6
Figura 4 Taxonomía del departamento de cadena de suministros. Elaboración propia, 2021. 11
Figura 5 Pasos del desarrollo del estado de cuestión. Elaboración propia, 2021. 12
Figura 6 Diagrama de flujo comprobación de fuentes. Elaboración propia, 2021 17
Figura 7 Diagrama de flujo selección de estudios. Elaboración propia, 2021 18
Figura 8 Búsqueda en Google Académico 20
Figura 9 Representación de resultados por Universidad. Elaboración propia, 2020. 23
Figura 10 Flujo prisma. Elaboración propia, 2020. 30
Figura 11 Nube de palabras. Elaboración propia, 2021 34
Figura 12 Mapa conceptual jerárquico del proyecto. Elaboración propia 2021. 35
Figura 13 Ejemplo de estructura en SC en empresa multinacional. 36
Figura 14 Representación de mariposa del riesgo de la cadena de suministro. Tomado de ManMohan S. Sodhi, Christopher S. Tang, 2012. 37
Figura 15 Estructura de una red perceptores multicapa (a) y una neurona artificial (b). Tomado de Moosmayer, 2013 42
Figura 16 Ontología del departamento de cadena de suministros. Fuente: Elaboración propia, 2021. 45
Figura 17 Arquitectura del sistema basado en ontología. Fuente: Elaboración propia, 2021. 45
Figura 18 Los tres ciclos de la ciencia de diseño. 49
Figura 19 Transformación de ideas abstractas a causas raíces. Fuente: Elaboración propia, 2021. 49
Figura 20 Organigrama de la población seleccionada. Fuente: Elaboración propia, 2021. 51
Figura 21 Mapa de empatía. Fuente: Adaptado de https://www.rdstation.com/es/blog/mapa-de-empatia, 2021. 53
Figura 22 Diagrama de flujo métodos de recolección y diagnóstico. Fuente: Elaboración propia, 2021. 55
Figura 23 Mapa de empatía. Fuente: Adoptado de https://www.rdstation.com/es/blog/mapa-de-empatia, 2021 56
Figura 24 Diagrama causa y efecto. Fuente: Elaboración propia, 2021. 57
Figura 25 Mapa de conceptos. Fuente: Elaboración propia, 2021. 58
Figura 26 Ventana de entrega representada en una distribución normal. Fuente: Guiffrida & Nagi, 2006. 60
Figura 27 Mapa flujo de valor “ciclo de la orden”. Fuente: Elaboración propia, 2021. 61
Figura 28 Diagrama SIPOC. Fuente: Elaboración propia, 2021 62
Figura 29 Ciclo de vida minería de datos (CRISP-DM). Fuente: IBM, 2021 63
Figura 30 Visualización de la ingesta de datos en la base de datos. Fuente: Elaboración propia, 2021. 68

Figura 31 ETL Carga de archivos Excel provenientes de ERP y almacén de datos. Fuente: Elaboración propia, 2021. 68
Figura 32 Carga de archivos excel a la base de datos. Fuente: Elaboración propia, 2021. 69
Figura 33 Contenedor de bucle para cada objeto. Fuente: Elaboración propia, 2021. 70
Figura 34 Flujo de datos: transformación y carga a la base de datos. Fuente: Elaboración propia, 2021. 70
Figura 35 Diagrama de flujo "Limpieza de datos". Fuente: Elaboración propia, 2021. 71
Figura 36 Inconsistencia en sistema ERP. Fuente: Sistema ERP 72
Figura 37 Matriz de gráficos de dispersión. Fuente: Elaboración propia, 2021. 77
Figura 38 Matriz de dispersión para variables "fuertes". Fuente: Elaboración propia, 2021. 78
Figura 39 Matriz de dispersión para variables “débiles". Fuente: Elaboración propia. 79
Figura 40 Matriz de gráficos de dispersión. Fuente: Elaboración propia, 2021. 80
Figura 41 Distribución de ítems con relación a su tiempo de entrega por mes. Fuente: Elaboración propia, 2021. 81
Figura 42 Reparto de los ítems según estado de entrega. Fuente: Elaboración propia, 2022. 82
Figura 43 Distribución de los tiempos de entrega según ciclo de vida del ítem. Fuente: Elaboración propia, 2021 84
Figura 44 Distribución del tiempo de entrega de fábrica. Fuente: Elaboración propia, 2021 86
Figura 45 Entrega de producto con soluciones adicionales. Fuente: Elaboración propia, 2021 87
Figura 46 Tiempo de recuperación de materiales por tipo de producto. Fuente: Elaboración propia, 2021. 88
Figura 47 Estatus de entrega por productos estrella. Fuente: Elaboración propia, 2021. 89
Figura 48 Distribución de los tiempos en Hub. Fuente: Elaboración propia, 2021. 90
Figura 49 Distribución de entrega según la prioridad del cliente. Fuente: Elaboración propia, 2021. 91
Figura 50 Distribución del estatus de entrega en productos incluidos en la planeación. 92
Figura 51 Resumen de los modelos utilizados en regresión logística. Fuente: Elaboración propia, 2021 94
Figura 52 Resultado del árbol de decisión en los 3 modelos. Fuente: Elaboración propia, 2021 96
Figura 53 Nuevo mapa flujo de valor “ciclo de la orden”. Fuente: Elaboración propia, 2021. 99

Abstract

Contexto: El uso de aprendizaje automático en terrenos de la cadena de suministro es un tema muy reciente y por su naturaleza debe de ser personalizado por tipo de industria y empresa.

Conocer por qué las entregas prometidas al cliente no cumplieron el tiempo pactado, es el nicho de comercialización de miles de directores y consultores externos. Una característica principal en este tema es la alta variabilidad de causas, ya que pueden ser influenciadas por estacionalidad, mejoras implementadas por el negocio, fenómenos esporádicos o mal planeamiento operativo tanto interna como externamente.

La idea de este proyecto fue estudiar las variables que, según juicio experto, explican o explicarán el atraso inminente de una entrega, resultado en un modelo predictivo que califica si un pedido será entregado a tiempo o no, de uso replicable para nuevos ingresos de pedidos.

Objetivo: Estimar el cumplimiento de la entrega de producto al cliente final para una empresa privada utilizando aprendizaje automático.

Método: Uso variado de metodologías: *Lean Six Sigma*, CRISP-DM y enfoque alternativo.

Resultado: Este tipo de tecnología fue usada por vez primera en el equipo de la empresa patrocinadora, el efecto causado fue la apertura de rigurosos proyectos paralelos de seguimiento y mitigación, además, de la futura implementación del modelo en ambientes de producción.

Conclusión: Este proyecto inició con un análisis retrospectivo que se convirtió en una herramienta predictiva con oportunidad de revolucionar la estimación en tiempos de entrega prometiendo fortalecer la relación cliente – empresa.

Palabras clave: cadena de suministros, riesgos, interrupciones, ciclo de vida de una orden, predicción, aprendizaje automático, métodos de clasificación, regresión logística, árbol de decisión.

Capítulo 1 Introducción

1.1 Generalidades

Este proyecto fue patrocinado por una corporación multinacional privada, cuyas políticas de confidencialidad fueron seguidas al pie de la letra; por lo tanto, se solicitó un congelamiento de 1 año para la publicación de esta tesis por parte de la Universidad Cenfotec. Por ello en cuanto a la eliminación de datos que puedan inducir el nombre de esta empresa, específicamente la sección 1.8, se utilizaron técnicas de maquillaje ético en los datos, además, no se publicaron estrategias claves de tipo operativas o administrativas de la corporación.

1.2 Antecedentes del problema

En los últimos años, la administración en la cadena de suministros ha cambiado sus estrategias en cuanto a la toma de decisiones. En este sentido, soportar iniciativas o problemas con análisis numéricos siempre ha sido un requerimiento fundamental para la aprobación o desaprobación de estas, con ello, se rescata la conciencia en cultura analítica empresarial del equipo; no obstante, la representación gráfica o el tiempo de obtención de estos datos no eran eficientes en tiempo ni precisión, al existir una posibilidad de error humano.

Para contrarrestar las necesidades del departamento, se creó una herramienta modalidad autoservicio que cubre temas de importancia para la cadena de suministro, incluyendo la funcionalidad de la evaluación sistemática del riesgo de entrega por orden; su lógica utiliza la metodología en estimación de tiempos de entrega estándar definido por el negocio años atrás, con esta, el resultado final es reevaluado por las administraciones de cada país, quienes determinan si es necesario sobrescribir la

fecha o categorías por juicio experto. Esta herramienta es el punto de partida para cualquier iniciativa y referencia a nivel mundial, cuenta con una excelente infraestructura tecnológica y un departamento de soporte dedicado.

La metodología en estimación de tiempos de entrega generaba falsas expectativas al usuario, pues ofrecía una fecha basada en tiempos estándar con altas dosis de positivismo al excluir dentro de su cálculo parámetros como: tiempos de atraso en fábrica o transporte, evaluación de la disponibilidad en materiales, periodo de aceptación en términos de facturación o disponibilidad de recibir el producto una vez construido por parte del cliente, omisión de días feriados, características adicionales al producto estándar, etc. Entre las consecuencias a este hecho se percibió el impacto directo en la rendición de cuentas para sus métricas a los directores regionales y también la alta dificultad para mantener la credibilidad del cliente por cambiar constantemente la fecha de entrega.

Tiempo después, la gerencia incluyó dentro de su estructura organizacional, personal con conocimiento en inteligencia de negocios agrupándolo en el departamento de estrategia y análisis de datos, quienes usando mencionada herramienta generan reportes en Microsoft Power BI. Es decir, se incrementó la visibilidad de datos especialmente de la evaluación del riesgo de entrega. Las características de estos reportes son: no inclusión de innovación tecnológica, exclusión de análisis de datos históricos (retrospectiva) y con alta probabilidad de ser replicado por cualquier usuario con conocimiento mínimo a Microsoft Excel y una conexión externa a base de datos, no hay oportunidad de sobreescritura o corrección en caso de error o anomalía hallada por

el negocio, las excepciones de resolución trascienden en promedio 8 horas (diferencia horaria).

A nivel operativo, el gerente actual originó un equipo llamado “x” conformado por los mismos miembros de su organización, cuyo enfoque es detectar las causas de atraso en una orden y crear acciones contrarestantes, sin embargo, su accionar no ha sido muy efectivo hasta hoy día.

1.3 Definición y descripción del problema

La alta gerencia del departamento de la cadena de suministros tenía una visibilidad parcial de la eficacia operativa traducida en el número órdenes despachadas a tiempo al cliente y su afectación monetaria en las ganancias; los resultados de la herramienta existente más el juicio experto descartaban parámetros alternos para catalogar anticipadamente si una orden sería entregada dentro de los tiempos contables (mes, cuarto o año fiscal). En otras palabras, si la región cumplirá o no sus metas de entrega a nivel compañía y satisfacción de la expectativa del cliente, la precisión de sus números se afinaban conforme se acercaba el cierre contable, no antes, sin dejar de lado, los problemas colaterales provocados por la señal errónea de entrega a los equipos logísticos y operativos, dejándoles un breve periodo de reacción con acciones contingentes.

Otro síntoma desentendido eran los cambios de números repentinos y sorpresivos, de un día a otro, millones descendían de la categoría probable a imposible de entregar; la razón del incumplimiento se atribuía a entes circunstanciales como por ejemplo Covid-19 en el 2021, clima, escasez de materiales, problemas en aeropuerto o aduanas, etc.

La herramienta existente es sobrescrita diariamente en un 90% por administradores de países pertenecientes a Latinoamérica, los datos ajustados manualmente del día de hoy, eran el resultado del día posterior, por causa de diferencia horaria entre los países y el tiempo consumido en la preparación de los informes. Si bien, al inicio de mes, esta diferencia no es notoria, a finales del mes o cuarto, la variación numérica causaba caos en los países afectados.

La oportunidad de mejora fue tomar todos estos tópicos, analizarlos y desarrollarlos con la ayuda de la tecnología analítica.

1.4 Justificación

-28%	**+1%**	**+4.5%**
Costo de proceso "Ajuste Manual"	Número de órdenes en riesgo de entrega que fueron recuperadas.	Mejora en resiliencia en términos de ganancias.

Figura 1 Beneficios mensuales de este proyecto. Elaboración propia, 2020.

Este proyecto abarcó problemas de la vida cotidiana laboral como la incertidumbre de la eficacia operativa, dificultad para calcular la resiliencia, personal enfocado en funciones diferentes a su rol y formulación o ejecución de proyectos que excluyen la ideología de la compañía; la figura 1 resume los beneficios cuantitativos de este proyecto.

En breve se explica el aporte del proyecto en cada rubro:

Incertidumbre de la eficacia operativa

Las métricas del equipo no eran fiables tanto en definición como en la data utilizada para el análisis, el alto mando constantemente solicitaba la justificación del aparente “incumplimiento” de los objetivos del departamento.

El proceso “Ajustes Manuales” consumía alrededor de 6 horas laborales diarias para cada uno de los responsables del departamento de la cadena de suministros en total, la responsabilidad recaía en 3 equipos: cadena de suministros, ventas y finanzas, el costo promedio de este proceso ascendía a 4 941 344 colones por mes. A propósito, el departamento de finanzas es quien confirmaba si la orden sería o no reconocida en el periodo fiscal. La figura 2 aclara la interacción entre dichos equipos.



Figura 2 Interacción entre equipos. Elaboración propia, 2020.

Dificultad para calcular la resiliencia

El desconocimiento numérico de los factores de interrupción usuales en el flujo de una orden entorpece los cálculos de la capacidad en resiliencia, al no contar con una base de comparación.

Personal enfocado en funciones diferentes a su rol

El personal calificado en roles de cadena de suministros se ocupaba el 95% de su tiempo laboral en la recolección de información, aclaración de discrepancias entre estados de las órdenes con los diferentes equipos (reuniones) para concluir su día, asignando las categorías para cada orden y para completar su informe, dejando en

segundo plano las acciones correctivas detectadas o toma de decisiones del negocio (0.5 horas libres). La figura 3 representa la distribución del tiempo empleado para cada tarea en un día laboral.

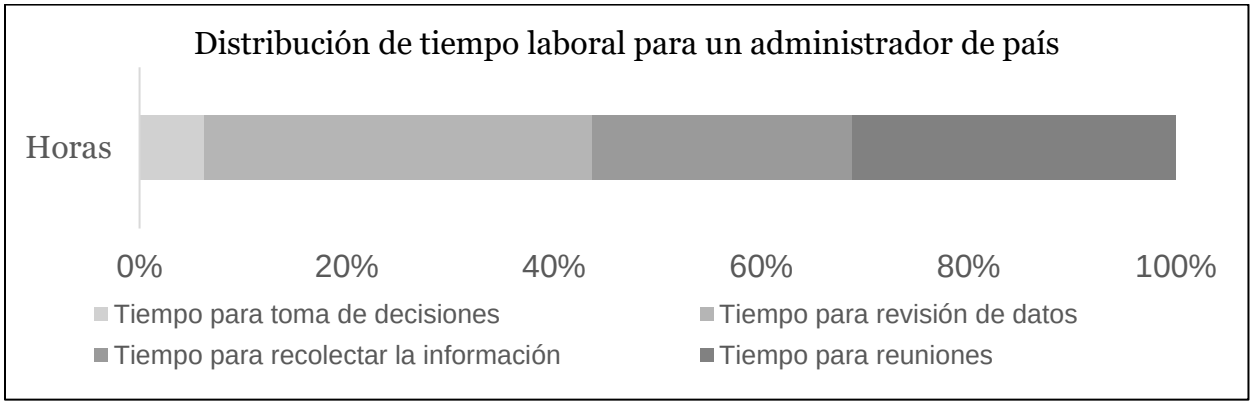


Figura 3 Distribución de tareas en un día laboral. Elaboración propia, 2021.

Los informes publicados por el departamento de estrategia y análisis de datos deben ser validados continuamente por los administradores del país para asegurar que los números representen la realidad de su operación.

Proyectos que excluyen la ideología de la compañía

Esta iniciativa se apega fielmente a la filosofía de la compañía:

- *“Falla y arregla rápido”*

Esta ideología va de la mano a los constantes cambios que un negocio enfrenta, lo que hoy es un “problema” o “tendencia”, mañana no lo es; por ende, el tema tratado fue tendencia en los últimos 6 meses y sostuvo un alto riesgo de ser descartado.

La premisa fue encontrar la causa de mayor peso para actuar ayer en las soluciones, que, dicho sea de paso, la solución administrativa u operativa no está dentro del alcance de este proyecto.

- *“Seamos curiosos”*

Contabilizar en días los posibles fallos en las entregas facilita las tareas de los administradores de país para mejorar su asertividad en su criterio, inclusive da paso a que cualquier persona de la compañía pueda realizar cálculos más aterrizados del tiempo de entrega de un producto.

“Somos una sola fuerza para el bien”

Los datos oportunos mejoran la comunicación y crean un ambiente de seguridad para la fluidez del bien común entre los departamentos de la compañía y satisfacción del cliente. *“La fuerza del bien nos acompaña” Star Wars.*

1.5 Viabilidad

A continuación, un análisis técnico, operativo y económico evidenciaron la viabilidad y factibilidad del proyecto.

1.5.1 Punto de vista técnico.

Se cuenta con estudios formales y experiencia en diversas áreas administrativas y técnicas, recientemente se obtuvo su maestría en tecnología de bases de datos. En la tabla 1, se encuentra el detalle de su conocimiento.

Tabla 1 Nivel de conocimiento de la investigadora

Conocimiento	Nivel de experiencia	Años	Nivel Académico
Ingeniería en Sistemas	●●●●●●●●○○	6	Maestría
Diseño Gráfico	●●●●●●○○○○	3	Técnico
Adm. de empresas	●●●●●●○○○○	2	Técnico
Lean Six Sigma	●●●○○○○○○○	2	Certificación Green Belt
Pensamiento del diseño	●●●○○○○○○○	1	Empírico

Fuente: Adaptado de la tesis Antony Calderón Vega, Universidad Cenfotec, 2020.

1.5.2 Punto de vista operativo.

Siguiendo la metodología de trabajo, se requirió, en cuanto a tiempo, de 5 colaboradores, 3 sesiones de 1 hora cada una.

El proveedor proveyó los siguientes requerimientos:

- Acceso a las bases de datos vinculadas con el proyecto.
- Acceso a servidor y espacio en el disco 10 GB.
- Uso de la licencia actual de SQL Server Integration Services (SSIS).
- Accesibilidad para instalar R Studio.

1.5.3 Punto de vista económico.

Todos los costos fueron asumidos por el patrocinador, se excluye el tiempo del investigador/profesional en informática.

La tabla 2 recopila el salario mensual de los roles participantes en el proyecto.

Tabla 2 Escala de salarios máximos de acuerdo con la página web tusalario.org de Costa Rica en colones, 2020.

Profesional	Salario Mínimo Mensual	Costo x hora
Gerente de suministro	₡ 2 643 481.24	₡ 11 014.50
Profesional Inteligencia Empresarial	₡ 1 725 088.24	₡ 7 187.87

Fuente: Elaboración propia, 2021.

La tabla 3 muestra el desglose del costo teórico del proyecto mensual para el patrocinador lo cual corresponde a ¢ 143 905.14. La fuente utilizada para el cálculo del precio fue la hoja de cálculo existente en la página oficial del proveedor.

Tabla 3 Recopilación de costos mensuales del proyecto para patrocinador, 2021.

Rubro	Horas	Costo	Total
Director (x1)	1	₡ 11 014.51	₡ 11 014.51
Profesional de almacén (x4)	4	₡ 3 222.66	₡ 12 890.63
Servidor incluye costo del uso de SSIS	30	₡ 4 000.00	₡ 120 000.00
Licencia R-Studio	30	₡0	₡0

Fuente: Elaboración propia, 2021.

Los costos mensuales para la investigadora oscilaron en un total de ₡287 517.70, ver tabla 4.

Tabla 4 Recopilación de costos mensuales del proyecto para investigador, 2021.

Rubro	Horas	Costo	Total
Profesional Inteligencia Empresarial	40	₡ 7 187.87	₡ 287 514.70

Fuente: Elaboración propia, 2021.

1.6 Objetivos

La Taxonomía de Bloom 1956, fue utilizada por su estructura jerárquica de objetivos que facilitan su redacción, además de ser el estándar de facto en la comunidad estudiantil costarricense. Los verbos fueron tomados de Fédorövv (2003)

1.6.1 Objetivo general.

Estimar el cumplimiento de la entrega de producto al cliente final para una empresa privada utilizando aprendizaje automático.

1.6.2 Objetivos específicos.

Conocimiento 1 Identificar las fuentes de datos que permitan la evaluación del riesgo o desvió de entregar a tiempo un pedido.

Comprensión 2 Distinguir los diferentes factores que intervienen en la entrega tardía de productos por parte del departamento de la cadena de suministros.

Aplicación 3 Desarrollar un modelo predictivo utilizando los factores encontrados en el objetivo 2.

Análisis 4 Desglosar el ciclo de producción de una orden a días calendario.

Sintetizar 5 Proponer una nueva metodología en estimación de tiempos de entrega, promoviendo un cambio cultural y mejorando la relación entre los clientes internos y externos.

1.7 Alcances y Limitaciones

1.7.1 Alcances.

A continuación, los entregables resultantes de este proyecto fueron:

- ✔ Prototipo del modelo predictivo en idioma inglés.
- ✔ Una demostración virtual del modelo en idioma inglés. La audiencia fue compuesta por el director y dos jefes del equipo del departamento de suministros de la empresa.
- ✔ Documentación del proyecto en idioma español.

1.7.2 Limitaciones.

El proyecto se enfocó en el área de cadena de suministros de una empresa internacional, para los países latinoamericanos cuyos pedidos y entrega de los productos al cliente final sucedieran durante el periodo: noviembre 2020 a mayo 2021.

No se incluyen las siguientes aristas del negocio:

X Análisis de tendencia de productos (compras, precios e inventarios).

X Indagación en la demanda o tendencia de clientes externos o internos, ni sus niveles de satisfacción con el servicio o producto.

X Evaluación de estrategias ecológicas.

X Estudio de costos operativos del departamento.

X Valoración operativa de tiendas o canales y depósitos aduaneros.

X Ciclo de vida de productos.

X Evaluación del proceso: compra de materia prima.

X Guía de acciones administrativas para corrección, ni de restauración de la operación en casos de detección de declives operativos.

La representación gráfica de los temas que fueron considerados en el proyecto se plasmó en figura 4, se enfocó en los procesos antes y durante de la producción.

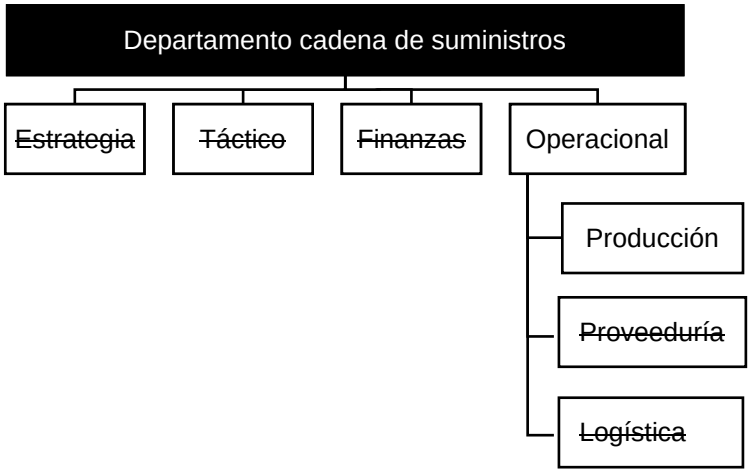


Figura 4 Taxonomía del departamento de cadena de suministros. Elaboración propia, 2021.

1.8 Marco de referencia organizacional y socioeconómico

Sección suprimida por solicitud de empresa patrocinadora.

1.9 Estado de la cuestión.

La revisión sistemática utilizada se basa en la plantilla de Biolchini, Gomes, Cruz y Horta (2005) por su descripción detallada de los pasos a seguir. Planificación, ejecución y resumen de los resultados son las etapas propuestas en procesos de ingeniería del software, ver figura 5.

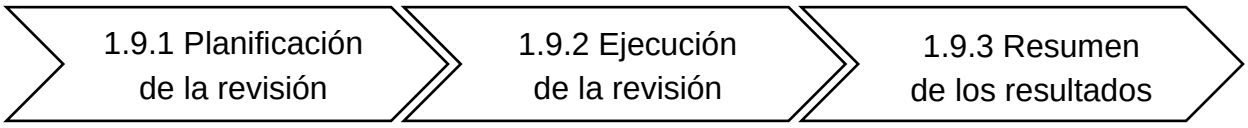


Figura 5 Pasos del desarrollo del estado de cuestión. Elaboración propia, 2021.

1.9.1 Planificación de la revisión.

En esta sesión se delimitó el tema a desarrollar y se construyó una guía a seguir en la búsqueda de material académico o científico.

1.9.2 Formulación de la pregunta.

Se dividió esta sección en dos partes iniciando con el foco de la pregunta (definición de los objetivos) para terminar con la amplitud y calidad de la pregunta (diseño de la pregunta).

1.9.2.1 Foco de la pregunta.

La redacción de la pregunta se enfocará en la investigación de metodologías o técnicas efectivas para la predicción de entregas tardías.

1.9.2.2 Amplitud y calidad de la pregunta.

Problema: La empresa enfrenta innumerables desafíos que alteran constantemente los planes administrativos definidos al inicio del mes o año, todo esto con miras a satisfacer la tendencia o necesidad de sus clientes actuales y los nuevos mercados a explorar; dichos planes se traducen en la disminución o aumento de tareas para los departamentos involucrados, la misión y lucha del departamento de cadena de suministros es la asertividad en la proyección de sus fechas de entrega y detención temprana de inconvenientes que no le permitan cumplir con las metas trazadas por la compañía.

a) Pregunta

La siguiente pregunta fue planteada: ¿Existen proyectos que tratan la detención o predicción de posibles entregas tardías?

b) Palabras claves y sinónimos

Se recopilaron palabras claves para realizar la búsqueda en motores de documentos académicos y científicos, hubo una inclinación a palabras del idioma inglés por su alto volumen en publicación.

La tabla 5, enumera la lista de palabras claves de búsqueda, las palabras en columna “Palabra” que contienen un “-” significa exclusión: palabras auxiliares que ayudan a delimitar el rango de búsqueda.

Tabla 5 Palabras clave

Palabra	Traducción en español	Acrónimos en inglés
Supply Chain	Cadena de suministros	SC
Machine Learning	Aprendizaje automático	ML
Prediction	Predicción	

Delivery delay	Entrega tardía	
Lead time delays	Retrasos en los plazos de entrega	
Shipment	Entrega	
Time risk	Riesgo de tiempo	
-Framework	Marco de Referencia	
-Finance	Finanzas	

Fuente: Elaboración propia, 2021.

c) Intervención: Se observó en la revisión sistemática los siguientes puntos:
   - Identificar el posible factor de innovación en el tema del proyecto.
   - Conocer las metodologías o técnicas más utilizadas para calcular y predecir riesgo de entrega de productos
   - Examinar las recomendaciones o limitantes de los autores que desarrollaron este tema

d) Control: No se cuenta con una base de información. Las palabras de búsqueda fueron el inicio de partida.

e) Efectos: Se esperaba encontrar nombres de metodologías o técnicas de predictibilidad.

f) Medida de salida: Las metodologías o indicadores de aplicación cuantitativa.

g) Población: Miembros del equipo de la cadena de suministros de los países latinoamericanos.

h) Aplicación: La información generada será de utilidad para toda empresa que posea su propio equipo de logística o cadena de suministros. Comunidad de administradores e ingenieros industriales.

i) Diseño experimental: No se aplicó un método estadístico, se analizó el contenido del documento versus la similitud al criterio buscado. La oportunidad de agregar material útil según criterio experto no fue negada.

1.9.3 Selección de fuentes.

En esta sesión se seleccionaron las fuentes de donde se obtuvo la materia prima de conocimiento para este proyecto.

1.9.3.1 Definición del criterio de selección de fuentes.

Las fuentes debían ser artículos, documentos, libros o conferencias en las áreas de cadena de suministros, ingeniería en sistemas o tecnología.

El criterio experto fue un argumento de gran valor y útil en el enrutamiento del tema.

Una limitante fue el acceso a la información, algunos documentos estaban sujetos a compra o inscripción a sitios web de origen desconocido.

1.9.3.2 Lenguaje de estudio.

El lenguaje universal inglés permitió la versatilidad de usar acrónimos conocidos en la industria, elevando la posibilidad de encontrar documentos, libros o conferencias que se alinearan con las palabras claves.

1.9.3.3 Identificación de fuentes.

En la tabla 6, se describe la planeación de la ejecución de búsqueda:

Tabla 6 Pasos de la identificación de fuentes.

Método de selección de fuentes ↓	Motores de búsqueda en la Web. Criterio de expertos dentro en la cadena de suministros.
Cadena de búsqueda ↓	Cadena#1 intitle:("prediction") intitle:("shipment") intitle:("Supply Chain" OR "SC") Cadena#2 intitle:("delivery delay") intitle:("SC" OR "Supply Chain") - intitle:("framework") + (ML or Prediction) Cadena#3 intitle:("supply chain") intitle:("time risk") + prediction Cadena#4 intitle:("lead time delays") intitle:("supply chain" OR "SC") –Finance
Lista de fuentes	Se utilizó Google Académico como motor de búsqueda, este contiene referencias a contenido científico y académico de varias casas publicadoras de investigación, además, de ofrecer funciones como la creación de un perfil configurable según los temas de interés.

Fuente: Elaboración propia, 2021.

Un punto rescatable de la tabla 6 fila cadena de búsqueda, se dividió en cuatro partes dado a la nula o poca información obtenida en las pruebas de búsqueda.

1.9.3.4 Selección de fuentes después de la evaluación.

Se realizó una revisión meticulosa de las revistas, verificando su nivel de prestigio, con ello, se aseguró que el material recopilado sea de buena fuente y aceptado por el área de estudio. En el caso de materiales publicados por universidades como tesis o investigaciones, se comprobó la posición de prestigio de la misma.

1.9.3.5 Comprobación de las fuentes.

La comprobación de las fuentes se basó en la utilización de un portal llamado Scimago, este fue diseñado para mostrar indicadores bibliométricos de una revista académica o científica y el manejo de rangos por país.

En la figura 6 se demuestra la ruta de comprobación de la fuente.

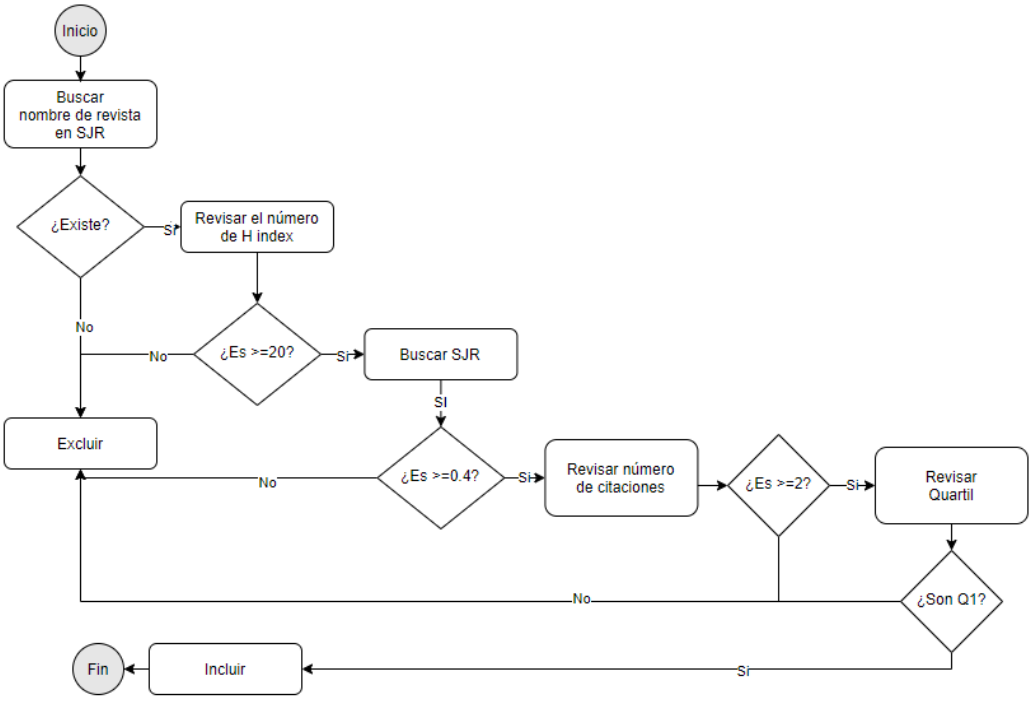


Figura 6 Diagrama de flujo comprobación de fuentes. Elaboración propia, 2021

En el caso de materiales de universidades, se evaluó a través de la página web topuniversities.com. Los parámetros evaluados fueron: número de posición de prestigio mundial, reputación académica y citaciones por facultad. Se tomaron en cuenta aquellas cuya posición de prestigio estén dentro de las primeras 50 universidades y citaciones >= 50.

Los documentos, libros o conferencias sugeridos por juicio experto, fueron excluidos de esta valoración.

1.9.4 Selección de los estudios.

En la figura 7 se describe, a través de un diagrama de flujo, la selección de estudios y su evaluación.

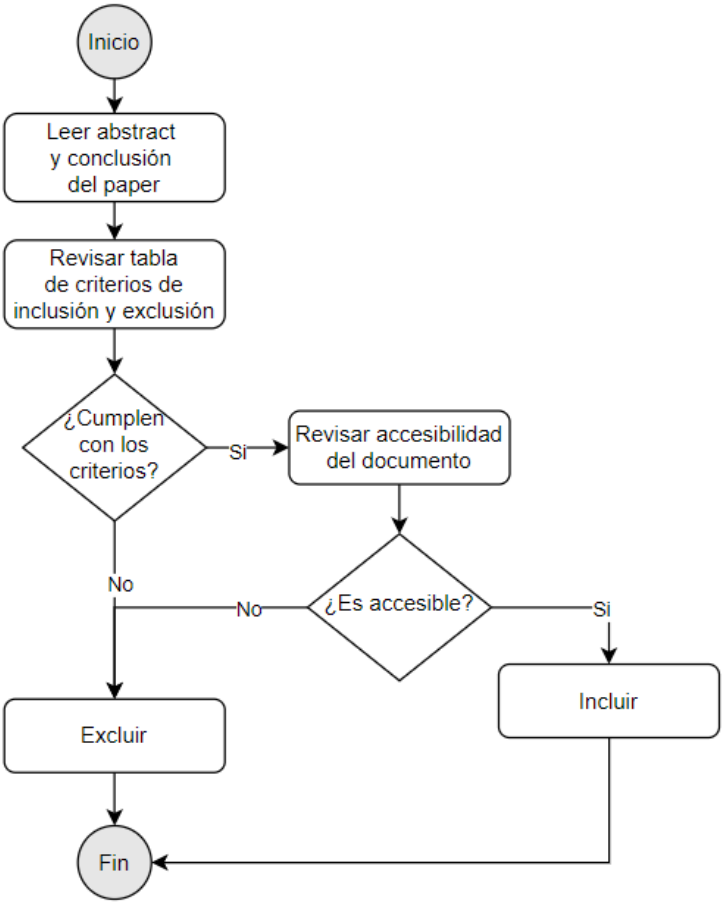


Figura 7 Diagrama de flujo selección de estudios. Elaboración propia, 2021

1.9.4.1 Definición del criterio de inclusión y exclusión de estudios.

En la Tabla 7 se describen los criterios para considerar o excluir los estudios encontrados.

Tabla 7 Criterio de inclusión y exclusión de estudios

Criterios de inclusión	Criterios de exclusión
Artículos, documentos, conferencias o libros que contengan aspectos directos con el tema en el resumen o *abstract* de su presentación.	Estudios empíricos
Posibilidad de casos implementados de éxito.	Artículos, documentos, conferencias o libros duplicados.
Artículos, documentos, conferencias o libros en idioma español o inglés.	Artículos, documentos, conferencias o libros con indicadores ajenos al estudio.
	Artículos, documentos o libros que no pertenecen a la industria de manufactura.

Fuente: Elaboración propia, 2021.

1.9.4.2 Definición de tipos de estudio.

No se limitó el tipo de estudio, solo que cumpliera con los criterios de inclusión y exclusión antes descritos.

1.9.4.3 Procedimiento para la selección de los estudios.

La selección de estudios tuvo una segunda etapa de aceptación después de ser evaluado por el punto 1.9.3.5 y fue la lectura ligera del contenido y gráficos del documento buscando la coincidencia con los siguientes tres puntos.

1. Contiene referencias bibliográficas que dan valor al proyecto.
2. Contienen la respuesta a la pregunta planteada en la sesión 1.9.2.2 punto a.
3. Contesta a las preguntas planteadas en la sesión 1.9.2.2 punto c.

Todo esto se aplicó también para aquellos documentos excluidos y de juicio experto, con la consigna de buscar una pista rescatable para el proyecto.

1.9.5 Ejecución de la revisión.

En esta sección se evidencia la ejecución de los puntos 1.9.1, 1.9.2 y 1.9.3.

1.9.5.1 Ejecución de la selección en la fuente: Google Académico.

Utilizando las cadenas de búsqueda del punto 1.9.2.2 b, se recopila un total de 9 resultados desglosados en la figura 8.



Figura 8 Búsqueda en Google Académico

1.9.5.2 Selección de estudios iniciales.

Tras realizar la búsqueda utilizando los parámetros mencionados se enumeran los resultados encontrados en la tabla 8.

Tabla 8 Lista de estudios a evaluar.

#	Referencia APA
1	Polim, R. (2016). Real-Time Supply Chain Analytics-Shipment Duration Prediction.
2	Kusrini, E., Sugito, E., Rahman, Z. M., Setiawan, T. N., & Hasibuan, R. P. (2021). Risk mitigation on product distribution and delay delivery: A case study in an Indonesian manufacturing company. In IOP Conference Series: Materials Science and Engineering (Vol. 722, No. 1, p. 012015). IOP Publishing.
3	Wang, T., Xing, Z., Hu, H., & Qu, X. (2019). Overbooking and delivery-delay-allowed strategies for container slot allocation. Transportation Research Part E: Logistics and Transportation Review, 122, 433-447.
4	Güller, M., Koc, E., Hegmanns, T., Henke, M., & Noche, B. (2015). A simulation-based decision support framework for real-time supply chain risk management. International Journal of Advanced Logistics, 4(1), 17-26.
5	Valero Berjaga, A. (2013). Development of a framework for the integration of real-time data in supply chain risk management using simulation tools (Bachelor's thesis, Universitat Politècnica de Catalunya).
6	Elangovan, D., Sundararaj, G., Devadasan, S. R., & Karuppuswamy, P. (2010). Analysis of time delays for loss reduction in an automobile components manufacturing organisation using a proactive supply chain risk management model. International Journal of Services, Economics and Management, 2(3-4), 385-411.
7	Mohamed, A. A. M., & Coutry, N. (2015). Analysis of lead time delays in supply chain: a case study. International Journal of Economics and Management Engineering, 9(6), 2065-2070.
8	Li, C., & Liu, S. (2013). A robust optimization approach to reduce the bullwhip effect of supply chains with vendor order placement lead time delays in an uncertain environment. Applied Mathematical Modelling, 37(3), 707-718.
9	Gudum, C. K. (2002). On the distribution of lead time delays in supply chains. Copenhagen Business School.

Fuente: Elaboración propia, 2021.

1.9.5.3 Evaluación de la calidad de los estudios.

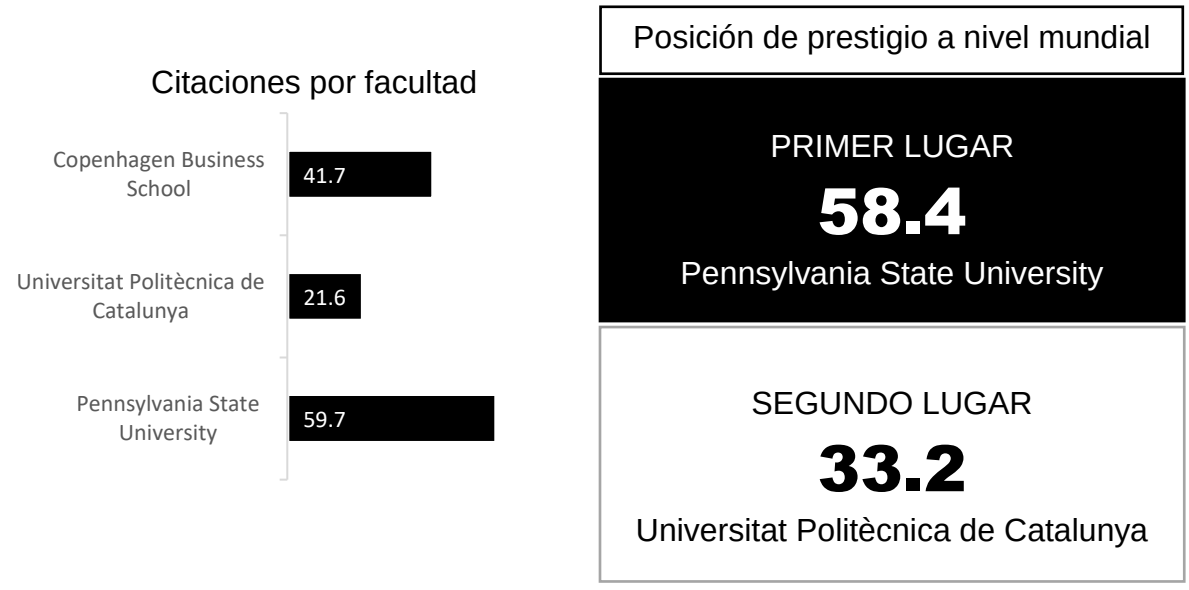
El proceso inició aplicando los criterios mencionados en el apartado 1.9.3.5, específicamente en el flujo de la figura 6; los resultados son recopilados en la tabla 9.

Tabla 9 Valoración de los estudios publicados por revista

#	Revista	H >=20	SJR > = 0.4	Citas > = 2	Q1	Incluido
3	Transportation Research Part E: Logistics and Transportation Review	103	2.3	7	Q1	SI
4	International Journal of Advanced Logistics	0	0	0	0	NO
6	International Journal of Services, Economics and Management	9	0.2	1	Q4	NO
7	International Journal of Economics and Management Engineering	0	0	0	0	NO
8	Applied Mathematical Modelling	103	1	4	Q1	SI

Fuente: Elaboración propia, 2021.

El paso final fue la evaluación de las universidades; la búsqueda realizada sugirió 3 instituciones. Es así como en la figura 9 se presenta un resumen de las calificaciones por documento: número 1 (Pennsylvania State), número 5 (Universitat Politècnica de Catalunya) y número 9 (Copenhagen Business School).



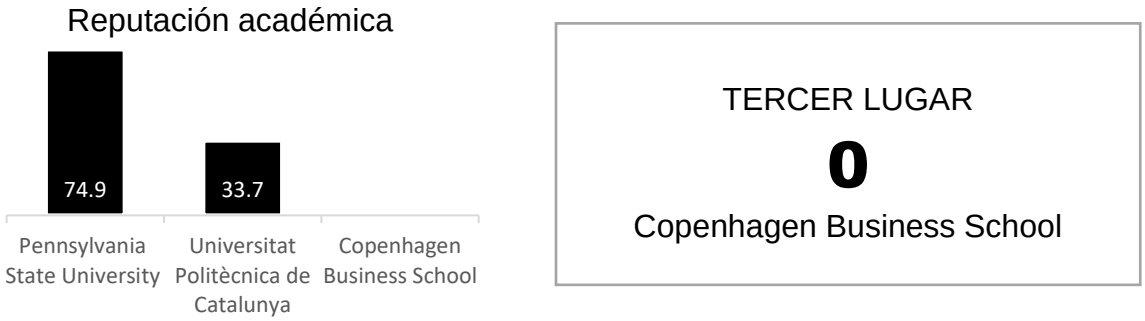


Figura 9 Representación de resultados por Universidad. Elaboración propia, 2020.

En la sesión Apéndices se encuentran las evidencias en imágenes de cada revista desde SJR y en el caso de las universidades topuniversities.com.

Seguidamente, en la tabla 10 se enumeraron los documentos, conferencias o libros incluidos en este proyecto por juicio experto o un valioso aporte enlistado en la bibliografía de los documentos antes evaluados.

Tabla 10 Documentos, conferencias o libros incluidos por juicio experto.

#	Tipo	Nombre
J-a	**Tesis**	Predicting Deviation in Supplier Lead Time and Truck Arrival Time Using Machine Learning *Link* https://odr.chalmers.se/bitstream/20.500.12380/257041/1/257041.pdf
J-b	**Tesis**	Supply chain visibility tools usage and their impact on on-time delivery: a case study of a fast-moving consumer goods small, medium enterprise in London *Link* https://bura.brunel.ac.uk/handle/2438/15663
J-b	**Artículo**	Improving supply chain visibility with artificial neural networks *Link* https://www.sciencedirect.com/science/article/pii/S2351978917305371
J-c	**Artículo**	Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., & McFarlane, D. (2021). Supply chain data analytics for predicting supplier disruptions: a case study in complex asset

		manufacturing. International Journal of Production Research, 58(11), 3330-3341. *Link* https://www.tandfonline.com/doi/abs/10.1080/00207543.2019.1685705

Fuente: Elaboración propia, 2021.

1.9.5.4 Revisión de la selección.

En la tabla 11 se aprecia la revisión de la selección y se sigue el flujo detallado en el apartado 1.9.4 Figura 7.

Tabla 11 Evaluación de criterios.

#	Documento	Criterios	Accesible	Incluido
1	Real-Time Supply Chain Analytics-Shipment Duration Prediction.	NO	SI	EX
3	Overbooking and delivery-delay-allowed strategies for container slot allocation. Transportation Research Part E: Logistics and Transportation Review, 122, 433-447.	NO	NO	EX
8	A robust optimization approach to reduce the bullwhip effect of supply chains with vendor order placement lead time delays in an uncertain environment. Applied Mathematical Modelling, 37(3), 707-718.	NO	SI	EX
J-a	Risk mitigation on product distribution and delay delivery: A case study in an Indonesian manufacturing company. In IOP Conference Series: Materials Science and Engineering (Vol. 722, No. 1, p. 012015). IOP Publishing.	SI	SI	IN
J-b	Predicting Deviation in Supplier Lead Time and Truck Arrival Time Using Machine Learning	SI	SI	IN
J-c	Supply chain visibility tools usage and their impact on on-time delivery: a case study of a fast-moving consumer goods small, medium enterprise in London	NO	SI	EX
J-d	Improving supply chain visibility with artificial neural networks	SI	SI	IN

Fuente: Elaboración propia, 2021.

Terminología: EX = Excluida IN = Incluida J= Juicio Experto

1.9.5.5 Extracción de la información.

A continuación, se documenta la forma en cómo se ejecuta y se extrae la información, conforme a los 4 documentos aprobados en las tablas 12, 13, 14 y 15.

Tabla 12 Extracción de información, documento J-a.

J-a	Nombre del estudio
Título	Risk mitigation on product distribution and delay delivery: A case study in an Indonesian manufacturing Company.
Descripción	Proyecto de identificación de riesgos de entrega para una empresa de manufactura.
Área	● Producción de manufactura ● Administración del riesgo ● Mitigación del riesgo ● Atrasos en entrega
Resumen	La mitigación del riesgo en una empresa de manufactura dependiente de terceros para la obtención de materiales o producción causa un riesgo que se torna en algunas ocasiones en una pesadilla por estar fuera de sus manos. Este documento propone una guía de administración del riesgo, resumidos en 4 pasos: 1. Identificación del riesgo (entendimiento) 2. Identificar consecuencias 3. Acciones reactivas o proactivas 4. Monitoreo de las acciones El estudio de 6 meses reveló la división de los riesgos originados entre el cliente y el proveedor. ● Proveedor: orden de compra, problemas de comunicación, capacidad de producción del proveedor, calidad del producto de los proveedores, documentos de envío, descarga y carga de mercancías por el transportista. ● Empresa: pronóstico de las compras de los clientes, planificación en producción, pedido de materias primas, control de calidad del material, capacidad de producción de la empresa, traslado de producto terminado al almacén y el proceso de recepción de documentos de mercancías.
Aspectos a destacar	Utilización de herramientas *Lean Six Sigma* para la documentación e identificación de los riesgos.

	La mayor cantidad de incidencia en los atrasos de entrega fue por capacidad del proveedor y calidad de sus productos, mientras que para la empresa las causas directas fueron pronóstico del pedido y el proceso de recepción de documentos de mercancías. Acciones sugeridas: transparencia, solicitud de colaboración en la planeación al proveedor, incrementar medidas de calidad, búsqueda de una solución automatizada para el pronóstico de compras y compartir el horario de entregas.

Fuente: Elaboración propia, 2021.

Tabla 13 Extracción de información, documento J-b.

J-b	Nombre del estudio
Título	Predicting Deviation in Supplier Lead Time and Truck Arrival Time Using Machine Learning
Descripción	Tesis aplicada a la empresa Volvo Group Service Market Logistics, su estudio se centraliza en los tiempos de ciclo del departamento de la cadena de suministro y los tiempos de entrega por los servicios de proveedores en logística (LSP siglas en inglés), dos modelos de aprendizaje automático fueron creados para revisar las condiciones y los posibles detonantes del atraso.
Área	● Aprendizaje automático ● Regresión y clasificación ● Tiempo de ciclo ● Cadena de suministros ● Logística
Resumen	El atraso en la disponibilidad de materiales repercute en los siguientes procesos. Los productos en obsolescencia reducen la existencia de materia prima e inventario. Las partes de uso primario tienen alto grado de prioridad en producción, las no críticas pueden causar atraso para completar el pedido. El precio de construcción de cada pieza puede ser un factor de atraso. Hay varios factores relacionados con los proveedores de materiales: priorización (basan su producción a clientes de pedidos mayoristas), escasez de mano de obra, avería de una máquina o volumen máximo de producción (la empresa no cuenta datos de proveedores, por lo que se basan en histórico de ventas para la creación del análisis). Por otro lado, el atraso por parte de logística está en gran parte en

	tiempos de espera de materiales, tiempo climático, político o económico. La información proveniente de la empresa no tiene el detalle suficiente para hacer un estudio completo de las posibles desviaciones, por ejemplo: disponibilidad de materiales.
Aspectos a destacar	Las técnicas de minería de datos utilizadas fueron: regresión y modelo de clasificación, regresión por motivos de generación de valor para la desviación del tiempo de entrega, mientras que la clasificación podría dar el resultado de tres clases, a saber, temprano, puntual y tarde (predictiva) respectivamente. Las sugerencias a la empresa destacan la mejora de las prácticas en la compra y la calidad de la información para perfeccionar los modelos propuestos.

Fuente: Elaboración propia, 2021.

Tabla 14 Extracción de información, documento J-c.

J-c	Nombre del estudio
Título	Supply chain data analytics for predicting supplier disruptions: a case study in complex asset manufacturing.
Descripción	En este documento, se discute la aplicación de análisis de datos para predecir interrupciones de la cadena de suministro utilizando datos históricos disponibles para un fabricante de equipos originales (OEM).
Área	● Cadena de suministros ● Aprendizaje automático ● Interrupciones ● Riesgos ● Análisis de datos ● Bosque Aleatorio ● Regresión Línea ● Regresión Logística
Resumen	Se realizaron tres fases: primero, una fase exploratoria para seleccionar y diseñar características potenciales que pueden actuar como predictores útiles de interrupciones, seguido por el desarrollo de una métrica de desempeño para calificar los métodos exitosos. En tercer lugar, se crea un diseño experimental para analizar sistemáticamente la tasa de éxito de diferentes algoritmos,

	parámetros algorítmicos, en el espacio de características seleccionadas.
Aspectos a destacar	Las interrupciones en el suministro pueden depender de una variedad de factores como problemas de calidad en una línea de producción, externos: como el tráfico, el clima, averías en las máquinas y, por lo tanto, su combinación puede presentar patrones aparentemente aleatorios cuando se utilizan sólo datos disponibles internamente. En cuanto al método predictivo el inconveniente fue la maldición de la dimensionalidad. Cuando el número de variables en el espacio de características es alto, los datos relacionados con cada atributo se vuelven escasos, lo que dificulta la significación estadística. Se trató de mitigar utilizando un valor de corte de cinco muestras de cada atributo durante la etapa de reprocesamiento; sin embargo, es necesaria una mayor experimentación.

Fuente: Elaboración propia, 2021.

Tabla 15 Extracción de información, documento J-d.

J-d	Nombre del estudio
Título	Improving Supply Chain Visibility with Artificial Neural Networks
Descripción	Para evaluar la eficacia del departamento de la cadena de suministros, se utilizaron redes neuronales artificiales para predecir la capacidad para cumplir con los pedidos entrantes y para anticipar qué nodos de la cadena de suministro recibirán un pedido para el próximo período a partir de datos históricos. Se cree que la información les daría a los gerentes la capacidad de ajustar los planes y minimizar los efectos que provienen de su incapacidad para satisfacer la demanda.
Área	● Inteligencia Artificial (AI) ● Red Neuronal Artificial (ANN) ● Cadena de Suministros (SC)
Resumen	El modelo está dedicado a un solo producto y consta de cuatro nodos: proveedor, producción, distribución y minorista. Los valores de estos parámetros fueron seleccionados junto con un conjunto de otros valores relacionados con los cuatro nodos: puntos de reorden, cantidades para ordenar, tiempos de transporte y capacidades de transporte para los cuatro nodos. Los nodos están conectados a través de un nodo de transporte que

	se establece con una capacidad y tiempo de transporte. Además, se supone que los almacenes tienen una capacidad de almacenamiento ilimitada y al comienzo la cadena de suministros tiene inventario capaz de alimentar al proveedor. En cuanto a las salidas, se registró el tiempo en el sistema de cada pedido y los niveles de inventario de cada escalón cada vez que ingresaba un nuevo pedido el sistema. Se aplicaron dos experimentos: ● En el experimento 1, la red neuronal artificial intenta predecir, durante un cierto periodo, si las próximas órdenes se cumplirán instantáneamente y si no, cuánto tiempo se tardará en terminarlas. ● En el experimento 2, la red trata de prever qué entidades llegarán a su punto de reorden y colocarán un nuevo pedido en sentido ascendente. En cuanto al experimento 1, los resultados fueron notables: la tasa de reconocimiento fue superior al 99,5% para predicciones para el período siguiente y más del 98% para las predicciones para los tres períodos siguientes. Estos resultados fueron comparados con un nuevo experimento, utilizando como entrada solo la información del nivel de inventario en el minorista. En ese experimento, la tasa de reconocimiento para períodos de tiempo más largos fue menor, lo que demostró que la visibilidad en las cadenas de suministro es importante y aumenta la precisión de la predicción. En el experimento 2, los resultados fueron 97% para las predicciones para el siguiente período y 75% para un horizonte temporal de tres períodos. A medida que se amplía el período de tiempo, la tasa de reconocimiento disminuye alrededor del 10-15% para cada período de tiempo que se agrega.
Aspectos a destacar	Las redes neuronales artificiales pueden ser utilizadas para predecir posibles interrupciones en las cadenas de suministro. Una recomendación de los autores fue la irrelevancia de extender los datos históricos, solo se necesitan el período actual y el anterior, los resultados son prácticamente los mismos.

Fuente: Elaboración propia, 2021.

1.9.6 Resumen de los resultados.

La figura 10 detalla el paso a paso del proceso de selección, utilizando el diagrama de prisma.

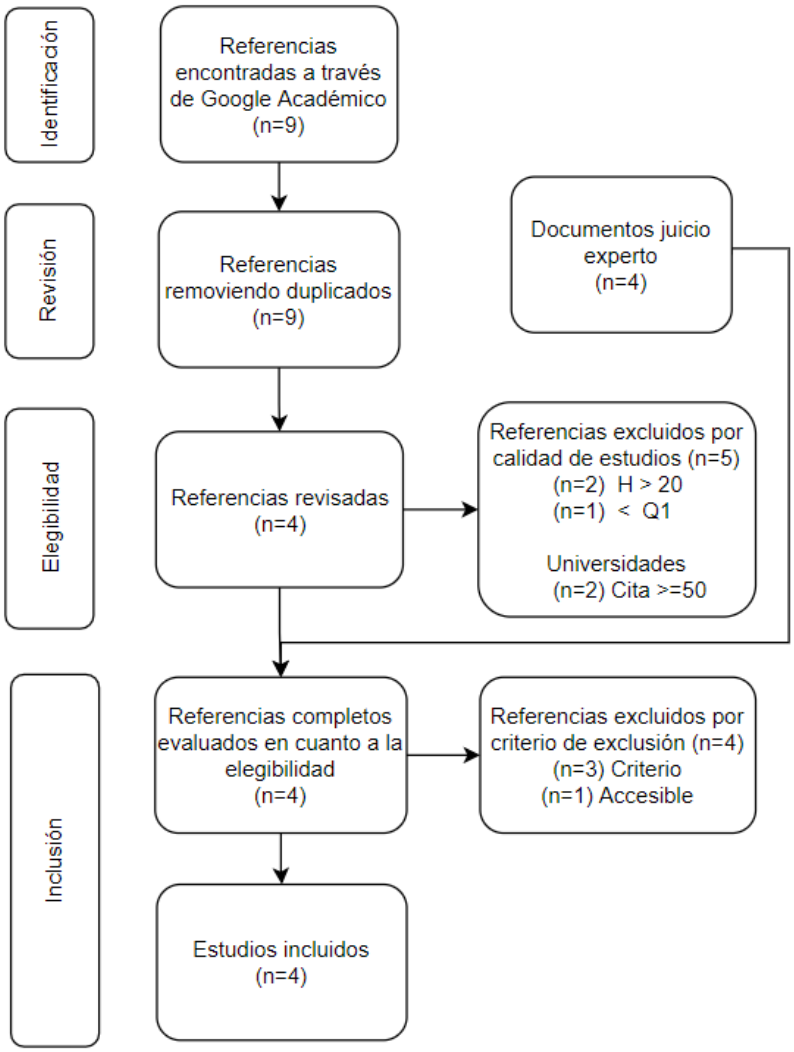


Figura 10 Flujo prisma. Elaboración propia, 2020.

Se sintetiza en la tabla 16 los parámetros textuales de referencia y la evaluación del nivel de acercamiento por tema incluido en el proyecto, se recopila en tabla 17 cada cuadrito equivale a un punto, si está una “X” = +1 en blanco = 0.

Tabla 16 Parámetros textuales de referencia para gráfico.

#	Documento
J-a	Risk mitigation on product distribution and delay delivery: A case study in an Indonesian manufacturing company. In IOP Conference Series: Materials Science and Engineering (Vol. 722, No. 1, p. 012015). IOP Publishing.
J-b	Predicting Deviation in Supplier Lead Time and Truck Arrival Time Using Machine Learning
J-c	Supply chain visibility tools usage and their impact on on-time delivery: a case study of a fast-moving consumer goods small, medium enterprise in London
J-d	Improving supply chain visibility with artificial neural networks
#	Pregunta
A	¿Existen proyectos que tratan la detención o predicción de posibles entregas tardías?
B	Identificar el posible factor de innovación en el tema del proyecto.
C	Conocer las metodologías o técnicas más utilizadas para calcular y predecir riesgo de entrega de productos
D	Examinar las recomendaciones o limitantes de los autores que desarrollaron este tema

Fuente: Elaboración propia, 2021.

Tabla 17 Evaluación del acercamiento del tema por pregunta

#Doc.	**A**			**B**			**C**			**D**			**Total**
J-a	X	X	X							X			4
J-b	X	X	X				X	X					5
J-c	X	X	X				X	X					5
J-d	X	X	X				X	X		X			6

Fuente: Elaboración propia, 2021.

Se concluye que el nivel de acercamiento de los documentos con el proyecto es de un 42%, por cuanto el estudio enfocado en redes neuronales alcanzó el mayor puntaje y como era esperado, el estudio de riesgos no tomó mucha importancia numérica del cual era de esperar.

Resumen de los hallazgos

La búsqueda de estudios en Google Académico fue un poco difícil, ya que los temas para dicho departamento son muy variados y personalizados, según el tipo de industria. Se encuentra mucha información sobre atrasos marítimos o aéreos, casos de planificación, minimización del efecto látigo y muy vagamente áreas de interés para este proyecto. Es por este motivo, se decidió segmentar la cadena de búsqueda en 4, para obtener resultados mínimos, puntuales, e inútiles. Alternativamente, se recurrió a la lectura de la bibliografía de los documentos, en estos se encontraron 2 de los artículos encasillados como juicio experto y se complementa el restante con lecturas de tesis al azar.

Los autores hicieron mucho énfasis sobre el obstáculo más grande: la calidad de los datos o la diversidad de estos. Las técnicas utilizadas se inclinaban por la reducción de productos o alcance en el número de proveedores, también recurrían a la justificante: “es posible realizar la predicción siempre y cuando se provean “X” y “Y” datos”.

El primer documento fue incorporado como una excepción porque no contaba con el factor predictivo buscado, pero el uso de herramientas en *Lean Six Sigma* daban significancia a los descubrimientos causales de atrasos en una empresa muy similar a la empresa patrocinadora.

Lastimosamente, no se encuentra un factor innovador y se interpreta el concepto de innovación como la creación de una idea desconocida por el mundo; sin embargo, al implementarse este proyecto a la empresa patrocinadora, se obtuvo una herramienta

innovadora para el departamento de cadena de suministros por la introducción de temas de predicción automática y utilización en vías de mejoramiento continuo.

Capítulo 2 Marco Conceptual

Basado en el contenido del estado de la cuestión se generó la siguiente nube de palabras con ayuda del portal Wordclouds.com (s.f.) (figura 11). Las palabras con mayor número de conteo son las que se describen en concepto y aplicación para este proyecto.



Figura 11 Nube de palabras. Elaboración propia, 2021

Con base en la nube de palabras, se crea el mapa conceptual jerárquico correspondiente, figura 12.

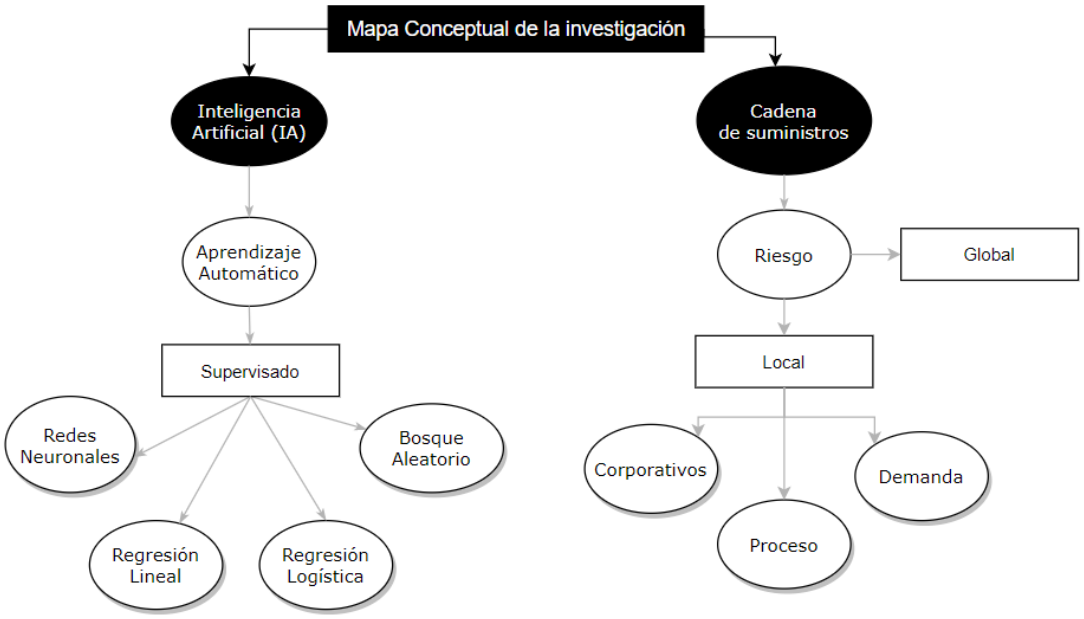


Figura 12 Mapa conceptual jerárquico del proyecto. Elaboración propia 2021.

2.1 Cadena de suministro (Supply Chain SC)

El diccionario de la Universidad de Oxford (s.f.) define a la cadena de suministro como: “La serie de procesos involucrados en la producción y suministro de bienes, desde que se fabrican, cultivan, etc., hasta que se compran o utilizan.”

En otras palabras, este se dedica operativamente en todas las actividades indispensables para que el producto o servicio llegue a su destino final, sea inventario o a las puertas del cliente. Su involucramiento en el negocio inicia cuando el departamento de ventas ingresa una orden de producción, el tratamiento de este orden depende de la estructura organizacional escogida por la compañía; la tendencia es dividir las tareas entre diferentes equipos como: procesamiento de la orden, planeamiento, manufactura, logística e inventario. La figura 13, ejemplifica una división de SC para una empresa.

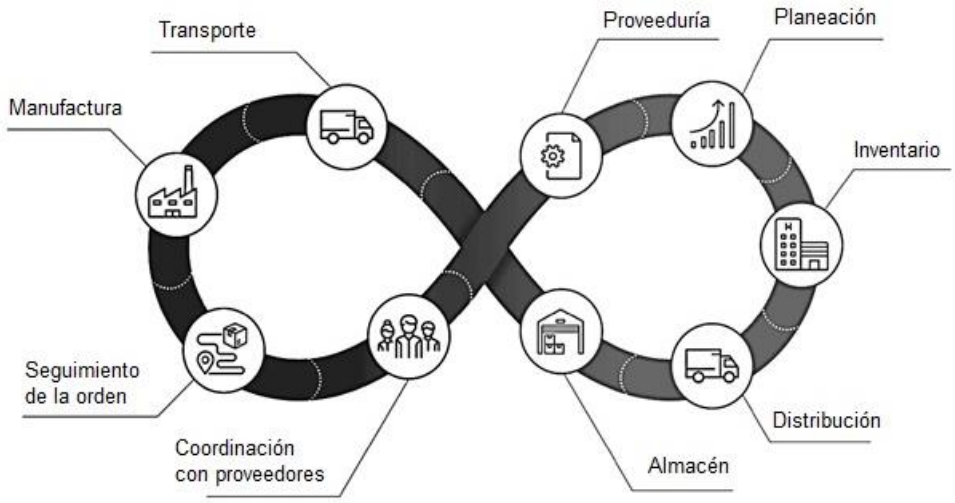


Figura 13 Ejemplo de estructura en SC en empresa multinacional.
Adaptado al español de https://www.fphighimpactpractices.org/briefs/supply-chain-management

2.1.1 Riesgo en cadena de suministros (Supply Chain Risk SCR).

Al respecto, COSO and WBCS (2018) indican que el riesgo es la probabilidad de que ocurra un evento e impacte en el éxito de la consecución de los objetivos estratégicos de la organización. La gestión de riesgos trata de identificar qué cosas están mal para recuperarse.

Asimismo, los riesgos no se pueden eliminar, se pueden aumentar o disminuir para perseguir una estrategia con la mínima inversión; sin embargo, estos se pueden explicar bajo dos mundos: globales y locales, en la figura 14 se graficó la mariposa del riesgo, donde se aprecia la interacción, entre ellos con la variable tiempo en una compañía.

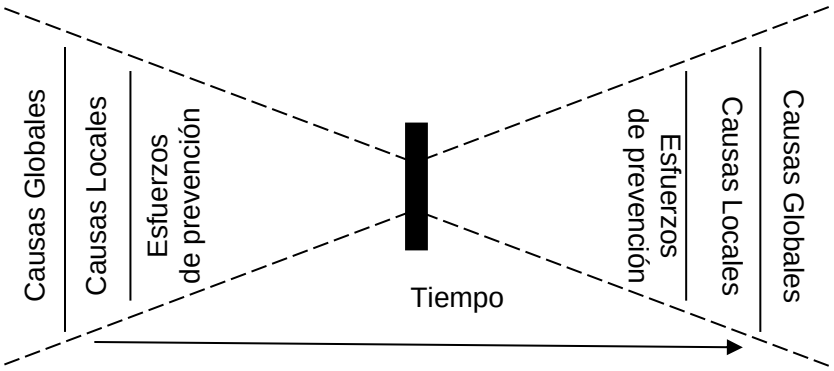


Figura 14 Representación de mariposa del riesgo de la cadena de suministro. Tomado de ManMohan S. Sodhi, Christopher S. Tang, 2012.

- Riesgos globales: ambiente exterior en donde opera la cadena de suministros, ejemplo: social, político, crisis crediticias o alzas de precio.
- Riesgos locales: comportamiento perjudicial dentro de la misma empresa o entidad: huelgas, bancarrota de nuestros aliados comerciales, contaminación/averías/obsolescencia en el proceso de producción, pérdida de los derechos de propiedad intelectual.

En este orden de ideas, Sodhi y Tang (2012) categorizan los riesgos por quién puede prevenir los eventos, por ende, recomiendan el diseño estructural en la cadena de suministro y los agrupan por procesos internos (manufactura, almacenamiento y seguimiento interno), tratamiento de la demanda (ventas) y compras. La segregación de riesgos se concretan en:

- Riesgos en la cadena de suministros: costos, fallos y compromiso del proveedor.
- Riegos en el proceso: inventario, capacidad y diseño.
- Riesgos en la demanda: Planeación de la demanda y cambios tecnológicos
- Riesgos corporativos: finanzas, visibilidad de operación, política y social, tipo de cambio, etc.

2.2 Inteligencia artificial (AI)

En este sentido, Legg y Hutter (2007) en su análisis, identifican dos características prerrequisito para considerar a un humano o una máquina inteligente:

1. La capacidad de elegir cuidadosamente sus acciones en una forma que conduce al éxito o al beneficio, en términos de algún tipo de objetivo o meta.
2. La capacidad de lidiar con un medio ambiente, pero con un abanico de posibilidades que no se pueden anticipar del todo.

2.2.1 Aprendizaje automático (Machine Learning ML).

Acerca de esto**,** el Diccionario Oxford (2002) indica que es:

> Una rama de la inteligencia artificial que se ocupa de la construcción de programas que aprenden de experiencia. El aprendizaje puede tomar muchas formas, que van desde aprender de ejemplos y aprender mediante analogía con el aprendizaje autónomo de conceptos y el aprendizaje por descubrimiento.

Asimismo, Barrios (2021) explica que son dos los tipos de aprendizaje, en donde, el aprendizaje supervisado (conceptos), la máquina tiene un “maestro” que le da a la máquina todas las respuestas. Es como decir que el maestro ya ha clasificado (etiquetado) los datos en gatos y perros, y la máquina está usando estos ejemplos para aprender y así clasificar con base en la predicción los casos futuros. En el aprendizaje no supervisado (descubrimiento): la máquina se queda sola con un montón de fotos de animales y una tarea para averiguar quién es quién. Los datos no están etiquetados, no hay profesor, la máquina está tratando de encontrar patrones por sí misma.

En el aprendizaje no supervisado hay dos tipos de tareas de este tipo:

- Clasificación: cuando se busca predecir una categoría, por ejemplo: pacientes enfermos y pacientes sanos.
- Regresión: cuando se busca predecir un valor numérico.

2.2.2 Regresión lineal.

Al respecto, Techopedia (2021) indica que la regresión lineal es un tipo de análisis estadístico que intenta mostrar una relación entre dos variables. Esta analiza varios puntos de datos y traza una línea de tendencia. También puede crear un modelo predictivo sobre datos aparentemente aleatorios, mostrando tendencias en los datos, como en los diagnósticos de cáncer o en los precios de las acciones.

Hay varias formas de calcular la regresión lineal. Uno de los más comunes es el método de mínimos cuadrados ordinarios, que estiman variables desconocidas en los datos, lo cual visualmente se convierte en la suma de las distancias verticales entre los puntos de datos y la línea de tendencia.

Los cálculos para realizar regresiones lineales pueden ser bastante complejos.

Afortunadamente, los modelos de regresión lineal se incluyen en la mayoría de los paquetes de cálculos principales, como Excel, R y MATLAB.

2.2.3 Regresión logística.

En este sentido, Techopedia (2021) afirma que la regresión logística es un tipo de análisis estadístico que se utiliza para predecir el resultado de una variable dependiente en función de observaciones anteriores. Por ejemplo, un algoritmo podría determinar el ganador de una elección presidencial en función de los resultados de elecciones anteriores y los datos económicos, un algoritmo más sofisticado también podría incorporar datos económicos y elecciones pasadas en su modelo. Otro algoritmo

podría intentar identificar qué usuarios de un sitio web harían clic en determinados anuncios. También se usa comúnmente en la preparación de bases de datos para clasificar datos para operaciones de extracción, transformación y carga (ETL).

2.2.4 Bosque aleatorio.

En un bosque aleatorio, los ingenieros construyen conjuntos de árboles de decisiones aleatorios para aislar más cuidadosamente el conocimiento de la minería de datos, con diferentes matrices de variables aplicadas. Una forma de describir la filosofía detrás del bosque aleatorio es que, dado que los árboles aleatorios se superponen, los ingenieros pueden crear sistemas para estudiar datos de forma redundante con los distintos árboles y buscar tendencias y patrones que respalden un resultado de datos determinado. Por ejemplo, si cinco árboles aleatorios proporcionan información sobre la misma variable de un subconjunto, y cuatro de ellos están de acuerdo, el algoritmo de aprendizaje automático puede utilizar ese "voto mayoritario" para construir modelos basados en probabilidades. En muchos tipos diferentes de aprendizaje automático, construcciones como el bosque aleatorio pueden ayudar a los sistemas tecnológicos a profundizar en los datos y proporcionar análisis más sofisticados.

Según González et al (2014) el bosque aleatorio es capaz de capturar las interacciones complejas con diferentes estructuras de datos y también es resistente a problemas de sobreajuste.

2.2.5 Redes neuronales artificiales (ANN).

Son modelos de inteligencia artificial basada en la estructura y capacidad de aprendizaje del cerebro humano (Moosmayer et al., 2013; (Rezaee et al., 2018).

El gran uso de las redes perceptoras multicapa MLP se debe a su idoneidad para afrontar problemas que involucran aproximación de funciones, clasificación de patrones, control de procesos y pronóstico de series de tiempo (Efendigil et al., 2009; Tkác y Verner, 2016). Las redes MLP pueden adaptarse a un entorno de uso específico por medio de un algoritmo de aprendizaje supervisado que utiliza un conjunto de muestras de las variables de entrada y salida. Se ha propuesto una gran diversidad de algoritmos de aprendizaje para entrenar redes MLP. El algoritmo de retropropagación (o descenso de gradiente) fue el primer aprendizaje método desarrollado para este propósito y ha sido adoptado en la mayoría de las aplicaciones empresariales (Tkác y Verner, 2016). Así, el uso de redes MLP permite que los sistemas de predicción de rendimiento se ajusten automáticamente al entorno de una determinada empresa mediante de datos históricos de rendimiento.

Como se muestra en la Fig.15 (a), la estructura de topología de una red MLP es formada por un conjunto de neuronas artificiales que están interconectadas por varias conexiones sinápticas. Está compuesto por una capa de entrada, al menos una capa intermedia y una capa de salida. Cada neurona de la capa de entrada representa una variable de entrada del modelo, mientras que las neuronas de la capa de salida denotan las variables de salida. Las neuronas de la capa oculta están encargadas de extraer las características de los datos y almacenarlos durante la etapa de formación de la red. Además, las neuronas de capa oculta también realizan la mayoría de las operaciones de procesamiento de datos en redes MLP (Akkoç, 2012; Tavana et al., 2016). Como se muestra en la Fig.15 (b), una neurona artificial está compuesta por

siete elementos: señales de entrada (x1, x2, … xn), sinápticas pesos (w1, w2, … wn), combinado lineal (Σ), umbral de activación (θ).

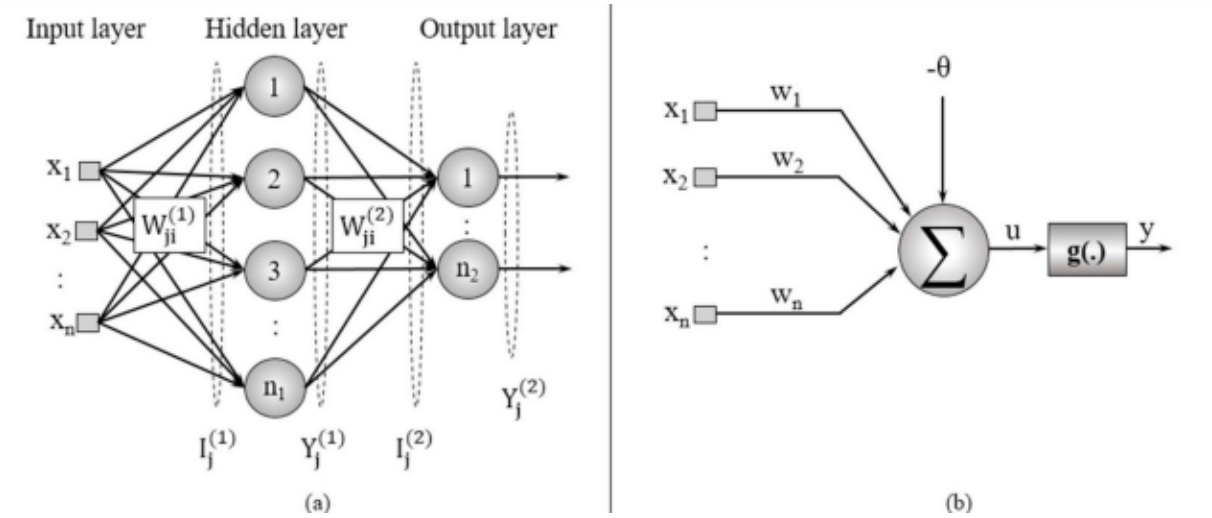


Figura 15 Estructura de una red perceptores multicapa (a) y una neurona artificial (b). Tomado de Moosmayer, 2013

Capítulo 3 Marco Metodológico

3.1 Tipo de investigación

Este proyecto fue la solución a un problema de un patrocinador, acorde con la definición de Universidad IBERO, las investigaciones aplicadas son aquellas en donde el conocimiento adquirido a lo largo del tiempo es invertido para resolver inconvenientes mediante el desarrollo de un nuevo producto, procedimiento o servicio.

3.2 Alcance investigativo

Este proyecto inicialmente tuvo una connotación exploratorio o formulativo y finaliza con rasgos de un proyecto descriptivo.

Los autores de “Técnicas de Investigación Social” engloban el propósito de una investigación exploratorio con objeto de… familiarizarnos con un tema desconocido, novedoso o escasamente estudiado. Son el punto de partida para estudios posteriores de mayor profundidad” (Ander - Egg, 1995).

Asimismo, Vásquez (2005) recopila en su estudio “Tipos de estudio y métodos de investigación”, la comprobación de este nivel respondiendo las siguientes preguntas:

1. ¿El estudio que propone tiene pocos antecedentes en cuanto a su modelo teórico o a su aplicación práctica?
2. ¿Nunca se han realizado otros estudios sobre el tema?
3. ¿Busca hacer una recopilación de tipo teórico por la ausencia del modelo específico referido a su problema de investigación? ¿Considera que su proyecto podría servir de base para la realización de nuevas investigaciones?

Al analizar las preguntas versus al tema de este proyecto, como se mencionó en el resumen de hallazgos del estado de la cuestión, no es un tema desconocido por el

mundo, pero sí lo era para el departamento en donde se ejecutaron los resultados de este proyecto.

Las interrupciones en la cadena de suministros es un tema predilecto para los ingenieros industriales, quienes buscan formas de erradicar el riesgo en compañías que pagarían millones al poder tener en sus manos la herramienta mágica que mantenga su integridad e identidad intacta ante los agente internos y externos; muchos de ellos se adentraron al tema de aprendizaje automático; sin embargo, encontraban muchas limitantes que por su desconocimiento en bases de datos o no presupuesto en sus proyectos, llegaban a un punto de reconocimiento del problema a medio término entonces documentaron estos obstáculos como fuera de su alcance, recomendaciones o mejoras al futuro.

La variedad de técnicas en este proyecto lo hacen ser un excelente material punto de partida y evolucionar en formas de seguimiento para una empresa con características similares al patrocinador.

Por otro lado, el involucramiento en la ejecución del proyecto redirecciona la investigación a un alcance descriptivo, pues, retomando palabras de (Ander-Egg, 2005) "…estos estudios describen la frecuencia y las características más importantes de un problema, para hacer estudios descriptivos hay que tener en cuenta dos elementos fundamentales; muestra e instrumento".

3.3 Enfoque

El enfoque alternativo fue abordado en este proyecto, no hace alusión a una interpretación cuantitativa, cualitativa o mixta; adicionalmente, se adecua a enfoques

investigativos para la informática con más facilidad. Naranjo (2019) fundamenta su implementación con base en tres dimensiones:

- Base ontológica: “Naturaleza del fenómeno a estudiar, ¿el fenómeno puede ser estudiado?” (Naranjo, 2019).

  La conceptualización del tema a investigar se puede apreciar en la figura 16 y 17, se describe al departamento y arquitectura conceptual de la solución.

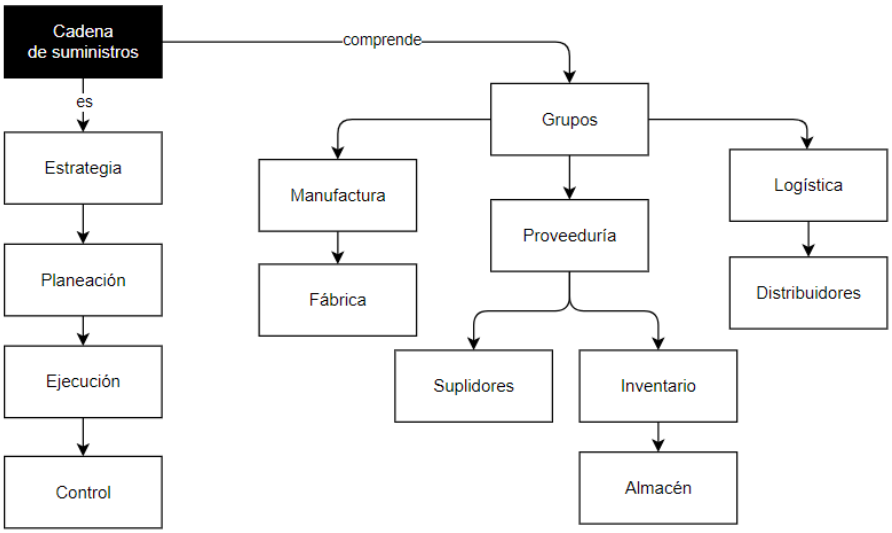


Figura 16 Ontología del departamento de cadena de suministros. Fuente: Elaboración propia, 2021.

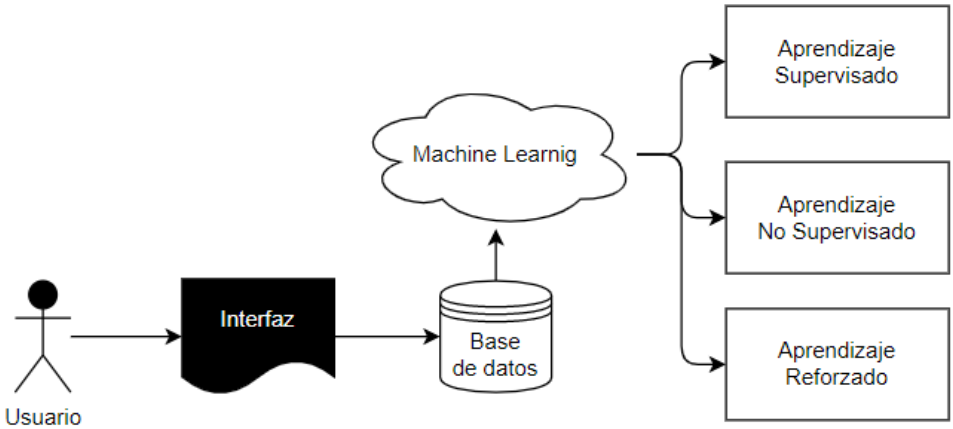


Figura 17 Arquitectura del sistema basado en ontología. Fuente: Elaboración propia, 2021.

- Base epistemológica: “Posición del investigador ante el fenómeno” (Naranjo, 2019). Para la investigadora un involucramiento total fue requerido para garantizar la efectividad de la implementación, iniciando por observación, experimentación, comprobación, finalizando con la respectiva documentación de resultados o hitos. No se pueden dejar de lado los usuarios, pues fueron parte fundamental para el éxito de este proyecto, su constante retroalimentación permitió una interpretación y transformación de los datos adecuados para el resultado esperado.
- Base axiológica: “Valores involucrados en el proceso investigativo, escala de valores o rúbrica que se usará para evaluar distintos aspectos” (Naranjo, 2019). La facilidad de implementación fue valorada a través del cuadro de evaluación sugerido en este enfoque; el resultado final (porcentaje) es contrapuesto a las siguientes consideraciones:

Tabla 18 Rúbrica para determinar el grado de dificultad de la implementación del proyecto.

Valor >= 80%	= Fácil de implementar
50 <= Valor > 80	= Complejo de implementar
Valor < 50	= Muy complejo de implementar

Rubro	Valor	Opciones		
Alternativas de implementación	20	Opción única: 5%	Dos opciones: 10%	Más de dos opciones: 20%
Tipo de algoritmo	30	Técnicas de IA: 10%	Técnicas convencionales con recursión: 15%	Código plano: 30%
Cantidad de líneas de código	25	Más de 2000 líneas: 15%	De 1000 a 2000 líneas: 25%	Menos de 1000 líneas: 35%

Preparación de los datos	5	3 tablas base o más: 5%	2 tablas base: 10%	1 tabla base: 15%

Fuente: Adoptada de Investigación en Informática: el enfoque alternativo, (Luis Naranjo-Zeledón, 2021)

La totalidad del valor es un 80% (tabla 18), es justo la señal positiva que se necesitó para proseguir con la implementación.

A continuación, la tabla 19 presenta los criterios de evaluación aplicables al modelo. Es indispensable el cálculo de la precisión, sensibilidad y especificidad.

Tabla 19 Rúbrica para determinar el grado de dificultad de la implementación del proyecto.

Evaluación del modelo	Criterio
Excelente	Precisión, sensibilidad y especificidad entre 100% a 75%
Aceptable	Precisión, sensibilidad y especificidad entre 74% a 50%
Malo	Precisión, sensibilidad y especificidad entre 49% a 0%

Fuente: Elaboración propia, 2021

3.4 Diseño

Este campo corresponde a las ciencias artificiales, por ende, la metodología de la ciencia de diseño encaja perfecto por su similitud al ciclo iterativo de la creación de diseño del software, cuyo valor agregado es la documentación de lecciones aprendidas, lo que llamaríamos una especie de bitácora.

Al respecto, Friedman, Maldonado y Bonsiepe y Simon (2017) argumentan que:

> Desde nuestro punto de vista, existe una diferencia de origen entre el diseño y la ciencia. La primera tiene por objetivo natural la generación de conocimiento y la comprensión de las leyes que rigen al mundo; la segunda en cambio tiene por objeto la generación de productos para satisfacer necesidades prácticas

> relacionadas con un usuario. …Mirar al diseño desde una sola óptica, es poseer una visión limitada de la disciplina.

En este sentido es agradable la idea de la fusión diseño y ciencia en campos como en la neurología, pues la distribución hemisférica del cerebro contradice el desarrollo del ser humano en ambas áreas, sin embargo, una solución a un problema se determina por el conocimiento y grandes dosis de la creatividad (crear, diseñar e innovar). “La metodología de ciencia de diseño plantea tres ciclos: el ciclo de relevancia, el ciclo de rigor y el ciclo de diseño” (Naranjo, 2019).

La herramienta principal para estos ciclos es la investigación y búsqueda de información útil para la construcción de un artefacto dentro de un contexto, ver figura 18.

- El ciclo de relevancia consiste en una revisión de las necesidades del mercado y del contexto en el que se utilizará el producto para entender mejor los requerimientos y requisitos del proyecto.
- El ciclo de rigor es una constante búsqueda de información relacionada con las necesidades identificadas, incluyendo soluciones anteriores a estas, es decir búsqueda de referencias y el conocimiento técnico necesario para llevarlas a cabo.
- En el ciclo de diseño se construyen las posibles soluciones al problema planteado y se evalúan por medio de distintas herramientas para validar su aporte (Robles, Vásquez y Naranjo, 2019).



Figura 18 Los tres ciclos de la ciencia de diseño.
Tomado de S. Robles-Sandoval, H. Vásquez-Carvajal, and L. Naranjo Zeledón.

3.4.1 Ciclo del diseño.

En este apartado se define la construcción y evaluación de artefactos.

El diseño del artefacto utilizó la grandiosidad de la metodología *Lean Six Sigma*; basados en la experiencia obtenida, la mencionada metodología provee técnicas para la recolección de información, convirtiendo ideas abstractas en causas raíces, ver figura 19.

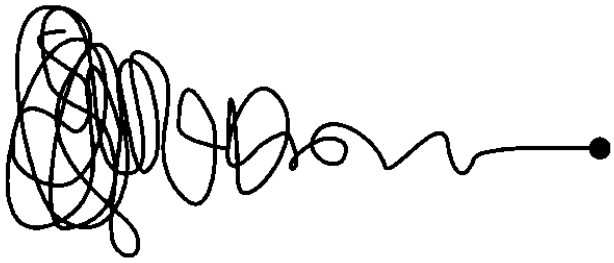


Figura 19 Transformación de ideas abstractas a causas raíces. Fuente: Elaboración propia, 2021.

Las etapas de esta metodología son: definir, medir, análisis, mejora y control, la tabla 20, nos describe en cada definición, que en conjunto se abrevian en inglés como DMAIC.

Tabla 20 Definición de las etapas DMAIC.

Definir	Medir	Análisis	Mejora	Control
En esta primera fase se definen los objetivos del proyecto, los miembros de los equipos de trabajo y las condiciones del problema.	Se debe evaluar el desempeño del proceso que se busca mejorar. Existen dos objetivos principales en esta fase: recoger datos que permitan cuantificar el problema e identificar las causas reales de este.	Se analiza la información sobre los resultados actuales e históricos. A partir de esto se puede descubrir la causa raíz del problema a enfrentar, así como establecer relaciones causa-efecto	A partir del análisis de las relaciones causa-efecto se realizan predicciones, las cuales, finalmente, conllevan a la generación de soluciones para el mejoramiento de los procesos.	Tras validar el funcionamiento de las soluciones, se procede a la implementación de controles que aseguren la continuidad del proceso.

Fuente: Elaboración propia, 2021

Las primeras tres etapas contaron con la participación de los interesados en conocer el porqué de las entregas tardías; la etapa control es parte de la propuesta de solución desarrollada en el capítulo 5 con ciertas modificaciones, ya que el objetivo consiste en identificar las causas no solucionadas, en el entendido que este proceso debe ser ejecutado y controlado por el departamento de operaciones de la empresa patrocinadora, es decir, este proyecto fue su punto de inicio en identificación y medición.

3.4.2 Ciclo de rigor.

La revisión sistemática de la base de conocimiento fue abarcada en la sección 1.9 estado de la cuestión de este documento.

3.4.3 Ciclo de relevancia.

¿Cómo se evalúa la eficiencia?, el departamento cuenta con cálculos optimistas como base, el resultado de este proyecto mostró alto grado de confianza por el proceso de descubrimiento y respaldo estadístico – matemático, también, la calidad de la documentación y bases bien sustentadas del origen de los mismos ganaron credibilidad e implementación inmediata en los métodos de cálculos futuros en la empresa patrocinadora.

3.4 Población y muestreo

La población para la recolección de datos no utilizará métodos de muestreo probabilísticos para dar validez a los hallazgos, la selección fue a convenir, se priorizan las necesidades de los jefes regionales de Latinoamérica en la cadena de suministros perteneciente a la empresa patrocinadora, más contactos estratégicos para la extracción de datos, más detalle en figura 20.

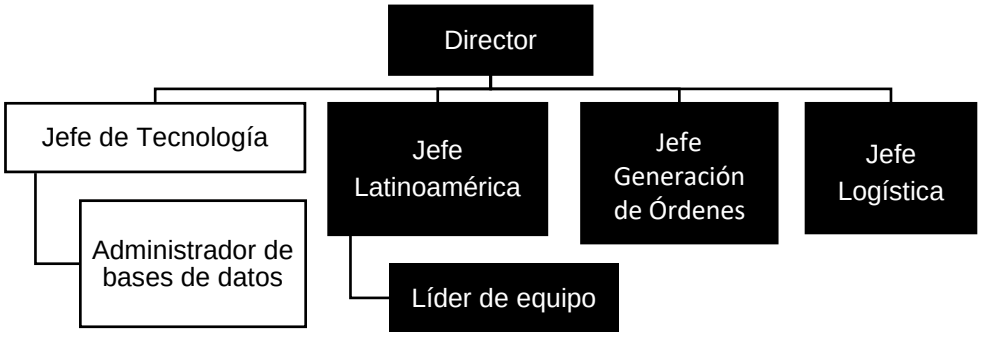


Figura 20 Organigrama de la población seleccionada. Fuente: Elaboración propia, 2021.

3.5 Instrumentos de recolección de datos

Los instrumentos de recolección de datos fueron tomados de la metodología *Lean Six Sigma*, técnicas usadas y comprobadas mundialmente.

El propósito de *Lean Six Sigma* es aumentar la capacidad y reducir el tiempo, variabilidad y defectos, fue desarrollado alrededor del año 1980 y rescata lo mejor de dos mundos *Lean y Six Sigma*, resultando una gran variedad de técnicas utilizadas por empresas de alto reconocimiento mundial como Toyota y Motola.

Muchas son las técnicas que se pueden aplicar para la recolección de datos, sin embargo, conociendo la disponibilidad del representante del negocio o cliente, se seleccionó la técnica de lluvia de ideas como base de inicio.

La lluvia de ideas es una actividad o método comercial en el que un grupo de personas se reúne para sugerir muchas ideas nuevas para un posible desarrollo (Diccionario de Cambridge, 2021).

Es así como el plan de acción pretendió reunir a un grupo focal con la idea de extraer sus requerimientos contemplando lo que entienden por su propia necesidad y la del director, además, se consideró la poca disponibilidad del equipo estratega en actividades extracurriculares a sus funciones. Una de las ventajas de este tipo de técnica, es la participación activa del grupo, lo cual le permite sentirse escuchado y crear ligamen de pertenencia en la discusión.

Por lo demás, el guion de trabajo para la lluvia de ideas fue el diagrama de empatía (figura 21). Esta es una herramienta que ayudó a formular las preguntas para la entrevista grupal; durante este tiempo, se buscó el enfoque del negocio y encontrar las dolencias que pueden convertirse en causas útiles de estudio. Al respecto, Pursell, (2021), directora de mercadeo en Hubstop, define en su blog personal al mapa de empatía como “una información detallada de cómo siente y cómo ve el mundo que le rodea el cliente con respeto a su problema.”

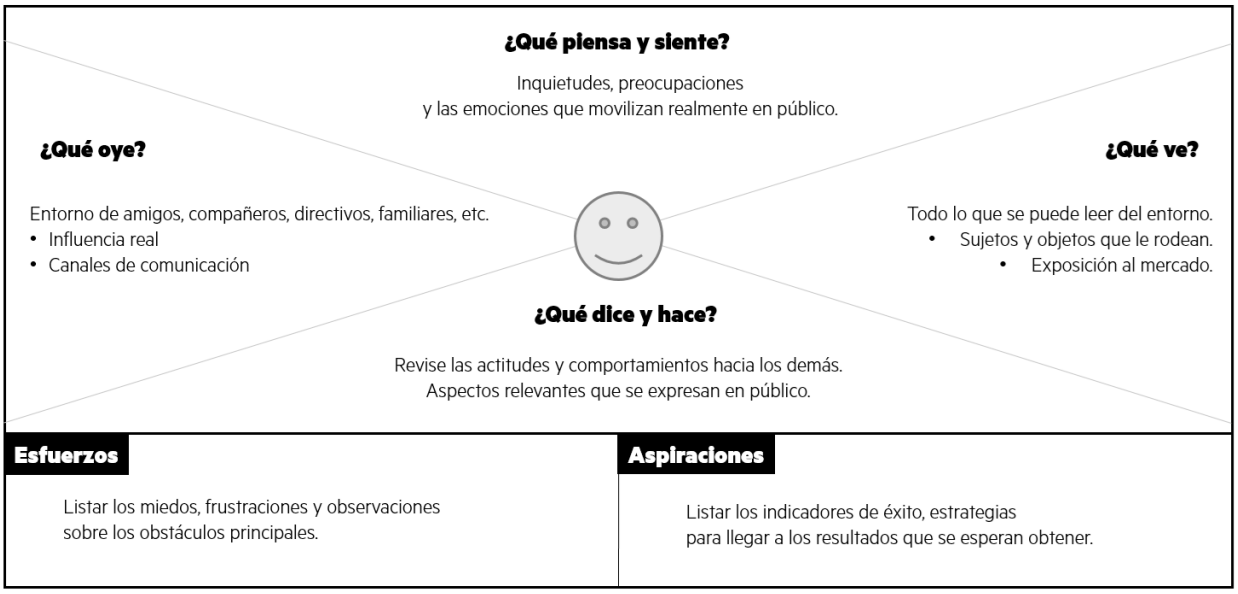


Figura 21 Mapa de empatía. Fuente: Adaptado de https://www.rdstation.com/es/blog/mapa-de-empatia, 2021.

Además, buscando la correcta documentación de los recursos digitales (datos, accesos) se recurrió al uso del gráfico SIPOC, en donde, para completar su construcción se indagó con las siguientes preguntas:

- ¿Cuáles son los departamentos dueños de los datos necesarios?
- ¿Cuáles son las fuentes que ofrecen estos departamentos?
- ¿Procesamiento de las fuentes?
- ¿Cuál es el resultado del procesamiento de las fuentes?
- ¿Quién usa ese resultado?

Vale indicar que Six Sigma Study Guide (2017) afirma que: “Un SIPOC es una herramienta de mapeo de procesos de alto nivel para ayudar a visualizar un proceso y sus influencias.”

3.6 Técnicas para análisis de información

Las técnicas para el análisis de la información seleccionadas fueron: diagrama Ishikawa (causa y efecto o diagrama espina de pescado), mapa conceptual y mapa de flujo de valor.

Definiciones de las técnicas de análisis de información

- Diagrama Ishikawa: representación gráfica del problema, sus causas y afectaciones. “En otras palabras, es una representación visual que se utiliza para descubrir la (s) causa (s) de un problema específico. El diseño de un diagrama se parece mucho al esqueleto de un pez. El diagrama de espina de pescado generalmente se dibuja de derecha a izquierda, con cada espina grande de pescado ramificándose para incluir espinas más pequeñas que contienen más detalles” (Six Sigma Study Guide, 2017).
- Mapa conceptual: Definición de conceptos. “Otra definición es que son instrumentos de aprendizaje fundamentados en las relaciones significativas que existen entre los conceptos.” (Universidad Estatal a Distancia, s.f.). Además, Hernández (2013) lo relaciona con “redes semánticas para organizar y representar el conocimiento.”
- Mapa de Flujo de Valor o Value Stream Mapping (en inglés), pretende graficar los procesos de la organización principalmente señalando los tiempos de ejecución y espera (Global Supply Chain and Operations Management, 2019). La herramienta conecta el proceso de flujo y análisis de costos visualizando la secuencia y haciendo aparentes los puntos de decisión "ocultos" marcando los tiempos de espera/gastos y sus recursos.

Capítulo 4 Análisis del diagnóstico

En este capítulo se describen los resultados de las técnicas de recolección y análisis de la información obtenida. En este orden, la metodología *Lean Six Sigma* apunta como primer paso conocer el punto de vista de nuestro cliente.

En la figura 22, se demuestra el resumen de la vinculación entre los métodos de recolección y diagnóstico.

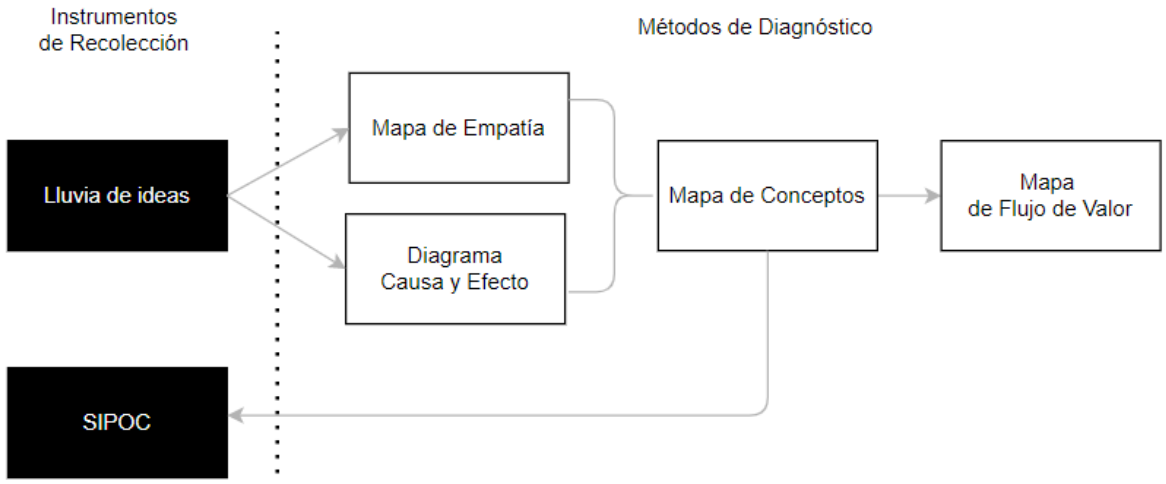


Figura 22 Diagrama de flujo métodos de recolección y diagnóstico. Fuente: Elaboración propia, 2021.

A continuación, se detalla la aplicación de cada método.

4.1 Técnica #1 Lluvia de ideas

Como se mencionó en las secciones 3.5 y 3.6, se aplicó la técnica lluvia de ideas, los participantes fueron el equipo de Latinoamérica, jefe generación de órdenes y jefe de logística, quienes dieron sus puntos de vista sobre una pregunta: ¿Cómo se puede detectar un posible atraso en el cumplimiento de la orden?, orientándola en 4 dimensiones:

a) ¿Cuál es su opinión?

b) ¿Qué ve en la compañía o externamente en relación con este tema?

c) ¿Qué escucha?

d) ¿Qué hace para mitigar el problema?

Las respuestas concedieron de participante a participante, se rescatan comentarios como: “los desvíos son secuenciales”; es decir, la orden no tiene todos los síntomas de atraso desde el día número uno en que es ingresada en los sistemas de la compañía, esta se puede quedar por días en una etapa, culmina y pasa a la siguiente etapa del ciclo, en esta puede o no quedar atorada por varios días más. Como nota adicional, los estatus de las órdenes son dinámicos y no son predictivos del siguiente obstáculo; sin embargo, existen códigos en los estatus de las líneas o información general que comprometen la entrega de la orden hasta que una o dos condiciones se cumplan. La figura 23 sintetiza la entrevista realizada a los participantes.

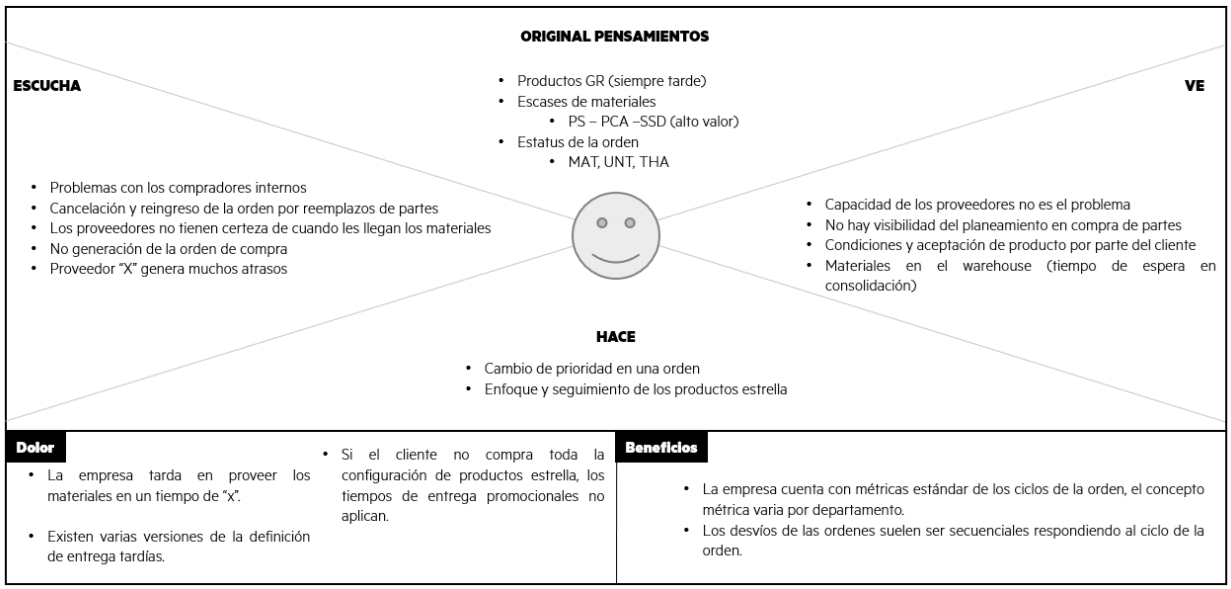


Figura 23 Mapa de empatía. Fuente: Adoptado de https://www.rdstation.com/es/blog/mapa-de-empatia, 2021

Las causas de desvíos no son desconocidas por la compañía, los temas de mayor peso rondan la escasez en materiales al igual que el desconocer si la orden está o no considerada dentro del planeamiento de compras mensual de la compañía. Este último es un factor determinante que puede incrementar hasta 30 días el tiempo de espera de una orden, figura 24 resume las principales razones de desvío.

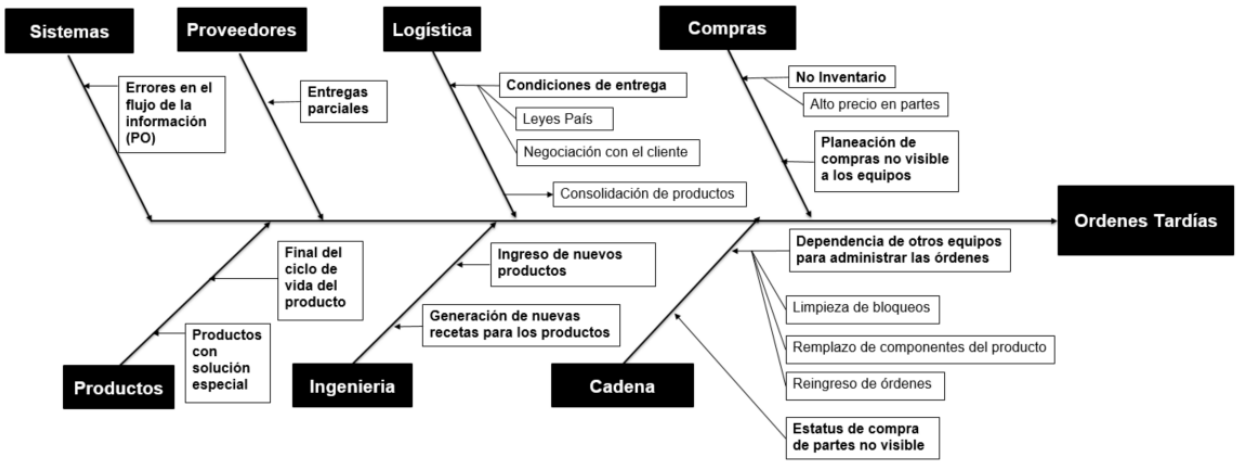


Figura 24 Diagrama causa y efecto. Fuente: Elaboración propia, 2021.

4.2 Técnica #4 Mapa conceptual

Las conclusiones obtenidas después de escuchar las opiniones de los diferentes expertos se concretaron en la figura 25.

En varias oportunidades los cuestionamientos/definiciones resultaron en respuestas múltiples como lo fue: ¿Cuál es su definición de una orden tardía?, curiosamente, existen varias definiciones dependiendo de la perspectiva cliente, proveedor o empresa.

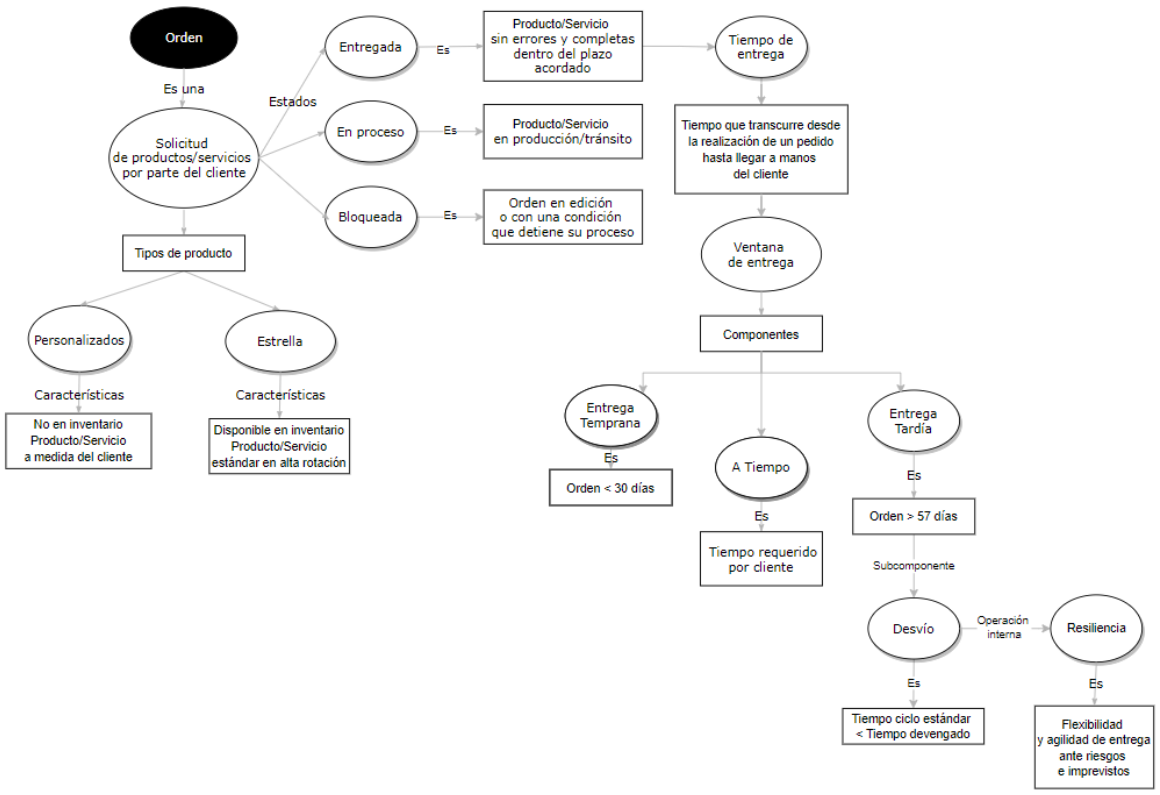


Figura 25 Mapa de conceptos. Fuente: Elaboración propia, 2021.

Complementando el trabajo realizado con el mapa de conceptos, seguidamente se extiende un poco más los conceptos generales:

Órdenes

Entregas de producto que se realizan sin errores y completas dentro del plazo acordado (Collins et al., 2001, p.7). Esto significa que el grado en que una empresa responde a sus clientes con un pedido preciso o completo que puede "hacer o deshacer" la relación cliente – proveedor.

Se dice que “Disponibilidad del producto (completitud del pedido y precisión)” es a menudo el factor más importante del servicio al cliente.

Tiempo de entrega

El tiempo de entrega de la cadena de suministro en un contexto logístico se define típicamente como el tiempo transcurrido entre el reconocimiento de la necesidad de realizar un pedido y la recepción de la mercancía (Blackstone y Cox, 2005).

El plazo de entrega también ha sido definido por otros estudios como el “Tiempo real promedio que transcurre desde la realización de un pedido hasta su entrega a un cliente” (Milgate, 2001, p.111). En este caso el tiempo de transporte no se toma en consideración, se asume que cuando un transporte sale de su punto de distribución es para la entrega directa al cliente.

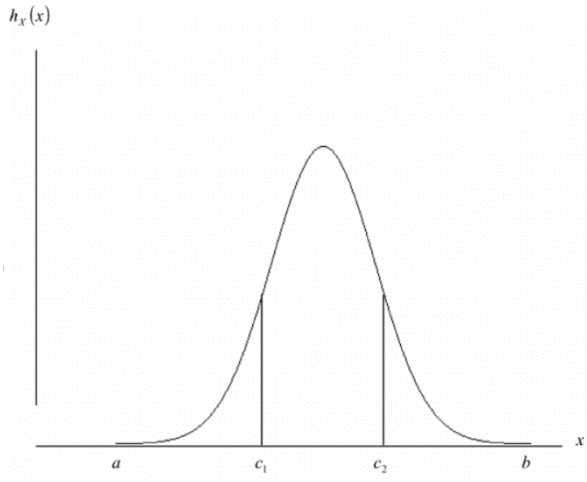
Ventana de entrega

Al respecto, Guiffrida & Nagi (2006) indican que una ventana de entrega se define como la diferencia entre las primeras fechas aceptables de entrega y la última fecha de entrega.

Los plazos de entrega son importantes para medir la entrega a tiempo, ya que la mayoría de los investigadores han señalado que este factor ayuda a las empresas a gestionar sus costos operacionales.

Los plazos de entrega permiten evaluar las desviaciones del plazo acordado para la entrega entre proveedores y clientes. Las desviaciones de la anticipación y el retraso de la fecha de entrega prevista son disruptivas para las cadenas de suministro, tanto temprano como tarde las entregas introducen a desperdicio, exceso de inventario y retención costos, retrasos en las entregas pueden conducir a la producción de paros y pérdida de la buena voluntad. Ambos factores pueden afectar la rentabilidad y

competitividad de una empresa, de ahí la importancia de mantener la entrega dentro los plazos especificados y acordados, en la figura 26 se grafican estos conceptos.



Simbología

a: plazo de entrega más temprano.

c1: comienzo de la entrega a tiempo.

c2: fin de la entrega a tiempo.

b: último plazo de entrega aceptable.

Figura 26 Ventana de entrega representada en una distribución normal. Fuente: Guiffrida & Nagi, 2006.

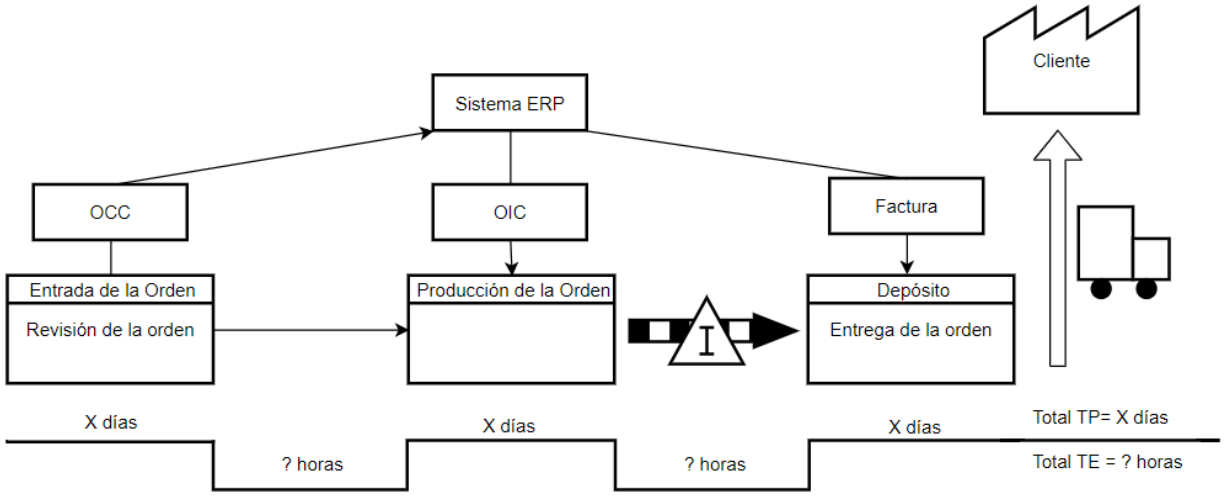
Resiliencia en la cadena de suministros

La flexibilidad de entrega se refiere al grado en que un proveedor/fábrica/cliente puede adaptarse a cambios en la fecha y hora de entrega acordadas. En otras palabras, el grado a la que se acomoda la incertidumbre en la entrega sin afectar al cliente o proveedor. Dicha flexibilidad representa las operaciones internas de la empresa proveedora: compras, manufactura y distribución mientras que la agilidad habla de la velocidad de respuesta.

Debido a la naturaleza de este proyecto, se utilizó el concepto de entrega tardía como la comparación de los tiempos de ciclo estándar versus el tiempo devengado en total.

4.3 Técnica #5 Mapa de flujo de valor

El mapa de flujo de valor, figura 27, muestra la hipótesis inicial en el tiempo de entrega de la cadena de suministro al cliente; en otras palabras, los tiempos del ciclo estándar. Es importante resaltar el desconocimiento de tiempos de espera en ciertos puntos del ciclo de la orden, representados con un signo de pregunta (?).



Tiempo de procesamiento (TP): Tiempo de transformación del producto que paga el cliente.
Tiempo de espera (TE): Tiempos de espera que no agregan valor.

Figura 27 Mapa flujo de valor "ciclo de la orden". Fuente: Elaboración propia, 2021.

Nota: Los valores representados con una "x" fueron suprimidos a solicitud del patrocinador.

4.4 Técnica #2 SIPOC

Los recursos digitales necesarios para este proyecto están almacenados bajo la infraestructura de "autoservicio". Los datos históricos necesarios eran inexistentes en la organización, pues, no eran de interés analítico, por lo que se requirió crearlos para poder desarrollar el proyecto. Se evidenciaron transformaciones importantes en datos tipo fecha, en donde se aplica determinada lógica para obtener los resultados

mixtos/alterados esperados por el negocio. En la figura 28, se muestran los actores a gran escala del proceso de datos en la corporación.

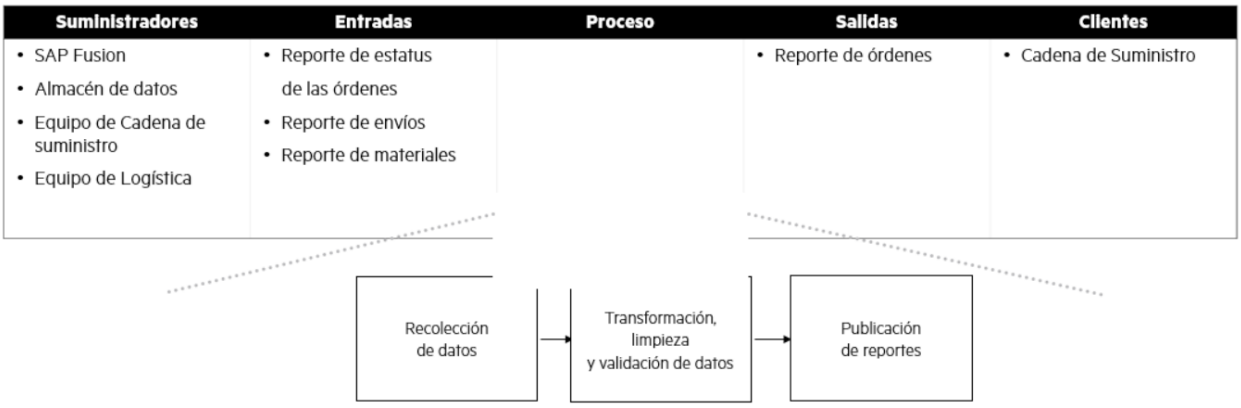


Figura 28 Diagrama SIPOC. Fuente: Elaboración propia, 2021

Capítulo 5 Propuesta de solución

En este capítulo se documentó la propuesta de solución que consta en dos partes:

1. Modelo predictivo del cumplimiento de entrega.
2. Propuesta de alteración a la metodología del cálculo de tiempos de entrega.

5.1 Modelo predictivo del cumplimiento de entrega.

La metodología utilizada fue CRISP-DM (*Cross-Industry Standard Process for Data Mining*), es una forma probada en la industria para guiar sus esfuerzos de minería de datos.

Como metodología, incluye descripciones de las fases típicas de un proyecto, las tareas involucradas en cada fase y una explicación de las relaciones entre estas tareas.

El modelo de ciclo de vida consta de seis fases con flechas que indican las dependencias más importantes y frecuentes entre las fases, Fig. 29. La secuencia de las fases no es estricta. De hecho, la mayoría de los proyectos avanzan y retroceden entre las fases según sea necesario (IBM).

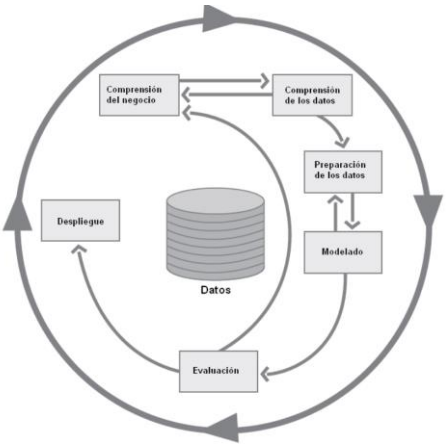


Figura 29 Ciclo de vida minería de datos (CRISP-DM). Fuente: IBM, 2021

A continuación, se procedió a desarrollar cada fase:

5.1.1 Entendimiento del negocio.

El departamento de la cadena de suministro contaba con un proceso manual para predecir si una orden cumplirá o no con el tiempo de entrega pactado con el cliente (Ajustes Manuales), empleando el 95% del tiempo laboral de los administradores de país divididos entre la recolección de información y aclaración de discrepancias con los diferentes equipos. En las sesiones 1.2, 1.3 y 1.4, se profundiza la situación en la cual la empresa patrocinadora se encontraba.

5.1.2 Comprensión de los datos.

Se trabajó con un conjunto de datos de 8755 observaciones y 81 variables para la orden de compra interna (OCI). Adicionalmente, 5212 observaciones y 62 variables para la orden de compra cliente (OCC). Se transformó en un set de datos con 5212 observaciones y 42 variables; es decir, se evaluaron 853 órdenes que obedecieron al criterio definido en las limitaciones 1.7.2.

En la tabla 21, se detallan las 42 variables utilizadas en este proyecto.

Tabla 21 Descripción de variables

Variable	Tipo	Concepto
Order	Número	Identificador único de la OCC.
Item	Número	Número del ítem de la OCC.
DelayProduction	Binario	0: La fábrica cumplió con el pedido según los tiempos de producción estándar. 1: La fábrica no cumplió con el tiempo de producción estándar.
DelayDeliverCustomer	Binario	0: Entrega total o parcial del producto al cliente a tiempo. 1: Entrega total o parcial del producto al cliente tarde.

Localization	Texto	Centro de distribución en país.
DeliveryCondition	Texto	Forma de entrega de producto: • Parcial: entrega por ítem. • Total: entrega por número de orden. • Cotización: entrega de órdenes varias que tienen el mismo número de cotización. Número generado por el equipo de ventas.
MonthCreated	Número	Mes calendario de creación de la OCC.
CustomerSegment	Texto	*Tipo de cliente, los clientes fueron divididos por su naturaleza: gobierno, privada, mixta, etc. Este campo cuenta con 12 tipos de clientes.
FinancialCustomer	Binario	0: El cliente no requirió un servicio financiero. 1: El cliente requirió un servicio financiero.
CustomerPriority	Texto	*Nivel de prioridad para el cliente. • Level1 = Prioridad 1. • Level2 = Prioridad 2. • Level3 = No hay prioridad.
ProductClass	Texto	*Clasificación por clase del producto, consta de 6 grupos.
Family	Texto	*Clasificación por familia del producto, consta de 5 grupos.
ProductType	Texto	Tipo de producto: • Intangible: servicio. • Stock: pieza de inventario. • *Group 1-5: productos clave. • Tangible: producto físico.
Ítems Núm.	Número	Número de ítems que componen una orden.
MixedOrder	Binario	0: OIC de un solo tipo de producto. 1: OIC con varios tipos de producto.
MultiSupplier	Binario	0: OIC construido por una sola fábrica. 1: OIC construido por varias fábricas.
Supplier	Texto	*Fábrica, son 10 en total.
CrossRegional	Binario	0: La construcción de producto requiere el uso de fábricas situada en un solo país. 1: La construcción de producto requiere el uso de fábricas situada en varios países.
AddSolution	Texto	0: OCC no requiere producto de solución especial 1: OCC requiere producto de solución especial.
StartProduct	Binario	0: El cliente no solicita un producto estrella.

		1: El cliente solicita un producto estrella.
w/oContract	Binario	0: El inicio de producción no requiere elaboración previa de un contrato. 1: El inicio de producción requiere elaboración de contrato previo.
Future	Binario	0: OCC para entrega en el cuatrimestre actual. 1: OCC para entrega en siguientes cuatrimestres.
Forecast	Binario	0: Material no previsto en planificación de compras. 1: Material previsto en planificación de compras.
SupplierPOCreationLate	Binario	0: Envío de la OIC a tiempo a fábrica. 1: Envío de la OIC tardía a fábrica.
CleaningTime	Número	Tiempo de limpieza de la OCC y OIC
WaitingTime1	Número	Tiempo de espera de la OCC entre el producto en centro de distribución país y la entrega final al cliente. No incluye tiempo de resolución.
WaitingTime2	Número	Tiempo de espera entre las diferentes etapas de la OIC iniciando desde la limpieza de la orden hasta finalización de la producción. No incluye tiempo de resolución.
MaterialWT	Número	Tiempo de espera para asignación de material.
HoldTime	Número	Tiempo de resolución de un incidente.
ProductionTime	Número	Tiempo de producción en fábrica.
TransitTime	Binario	Tiempo de entrega de la fábrica al país de distribución.
DeliveryTime	Binario	Tiempo total de entrega parcial o total de la OCC.
EscalacionFlag	Binario	0: OIC no fue escalada a fábrica. 1: OIC fue escalada a fábrica.
RedireccionFlag	Binario	0: OIC fue redireccionada a otra fábrica. 1: OIC no fue redireccionada a otra fábrica.
TestFlag	Binario	0: OIC no sufrió problemas en pruebas. 1: OIC tuvo problemas en pruebas.
Hold1Flag	Binario	0: OIC no presentó incidentes tipo 1 1: OIC presentó incidentes tipo 1
Hold2Flag	Binario	0: OIC no presentó incidentes tipo 2 1: OIC presentó incidentes tipo 2
Hold3Flag	Binario	0: OIC no presentó incidentes tipo 3 1: OIC presentó incidentes tipo 3
Hold4Flag	Binario	0: OIC no presentó incidentes tipo 4 1: OIC presentó incidentes tipo 4
Hold5Flag	Binario	0: OIC no presentó incidentes tipo 5

		1: OIC presentó incidentes tipo 5
Hold6Flag	Binario	0: OIC no presentó incidentes tipo 6 1: OIC presentó incidentes tipo 6
Hold7Flag	Binario	0: OIC no presentó incidentes tipo 7 1: OIC presentó incidentes tipo 7

Fuente: Elaboración propia, 2021.

Nota: (*) Se omite la descripción/clasificación por solicitud del patrocinador.

5.1.3 Preparación de los datos.

Se inició con la recolección de datos a partir de la lluvia de ideas creada en la figura 23 utilizando las entradas indicadas en la técnica SIPOC Fig. 28.

5.1.3.1 Extracción y carga de datos.

Los datos fueron extraídos manual y masivamente desde el sistema ERP y recursos alternos, consolidando la información necesaria para cumplir con los objetivos planteados en el capítulo 1. La tabla 22 específica el objeto/aplicación utilizado.

Tabla 22 Detalle de las variables usadas por recurso

Recursos	Tablas	Objetivo
SAP Fusión	VBAP	Información general de OCC y OIC
	VBAK	Información sobre los ítems de una orden Verificación de líneas canceladas = motivo de rechazo = “x”
	JCSD	Información de los estatus de OCC y OIC
	VBFA	Información de entrega OCC y OIC
	MB51	Inventario de productos entregados OCC
Almacén de datos	Órdenes app	Se extrajo toda la información relacionada a las variables en la tabla 21.
Cadena de suministro	Base de datos del equipo	Comentarios del estatus OIC.

Fuente: Elaboración propia, 2021.

El resultado de las extracciones fueron archivos Excel, los cuales fueron posteriormente ingeridos a la base de datos a través de la utilización de ETL (Extracción, transformación y carga), la figura 30 ilustra el proceso mencionado.

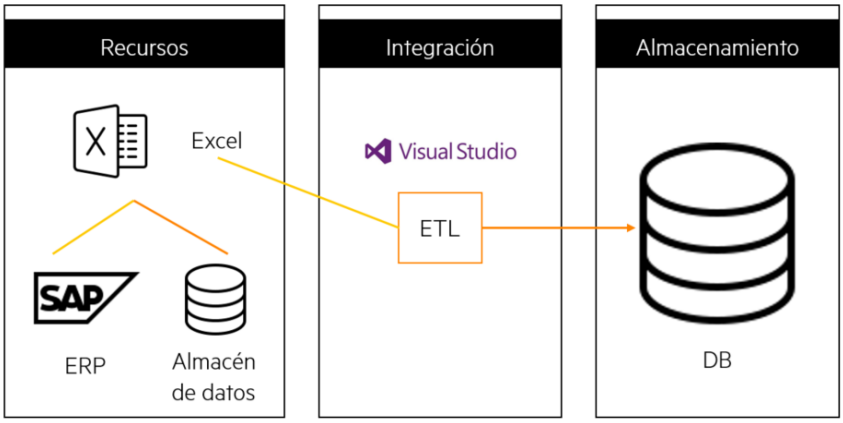


Figura 30 Visualización de la ingesta de datos en la base de datos. Fuente: Elaboración propia, 2021.

Un único ETL con dos paquetes se encargaron de llenar las tablas de la base de datos; la figura 31 muestra el flujo de los datos para cada uno de los archivos. Ninguna de las tablas son dependientes entre sí, por lo tanto el orden de la carga a la base de datos no es relevante.

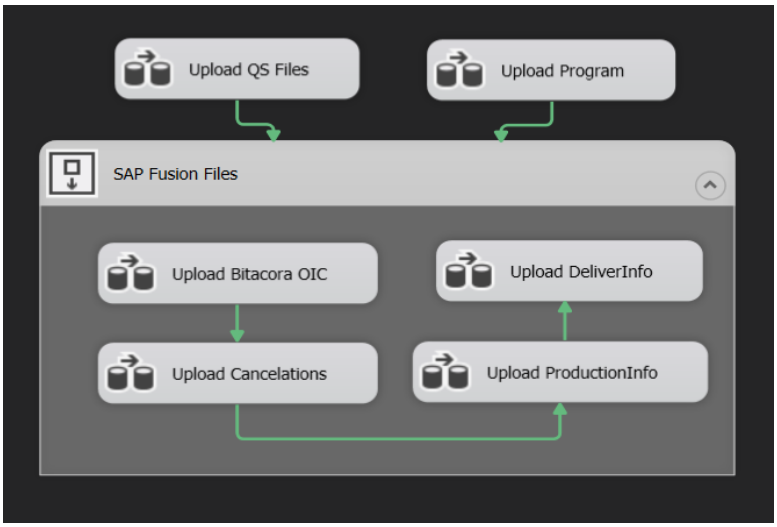


Figura 31 ETL Carga de archivos Excel provenientes de ERP y almacén de datos. Fuente: Elaboración propia, 2021.

Los 8 controles de flujo (*Upload QS Files, Upload Program, Upload Bitacora OIC, Upload Cancelations, Upload DeliveryInfo and Upload ProductionInfo*) manejan los mismos elementos para el flujo de datos, la figura 32 muestra los pasos.

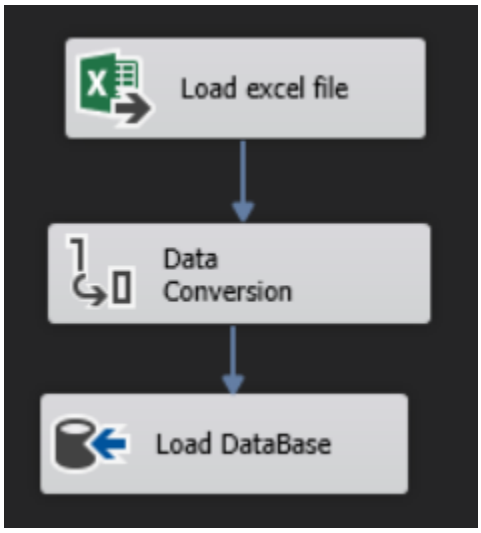


Figura 32 Carga de archivos excel a la base de datos. Fuente: Elaboración propia, 2021.

Para el caso del reporte de estatus de las órdenes, se desarrolló un ETL por aparte, se ocupaban cargar 73 archivos Excel. Se utilizó un objeto llamado “contenedor de bucle para cada objeto”, con esto se definió el lugar de almacenamiento de los archivos; seguidamente las conversiones de tipo de dato demandada por la base de datos, obtención del nombre del archivo con la finalidad de saber cuál fue el día de actualización del comentario para finalmente almacenar los comentarios a la base de datos. Estos pasos son ejecutados por el contenedor para cada uno de los archivos automáticamente. Ver figura 33 y 34.

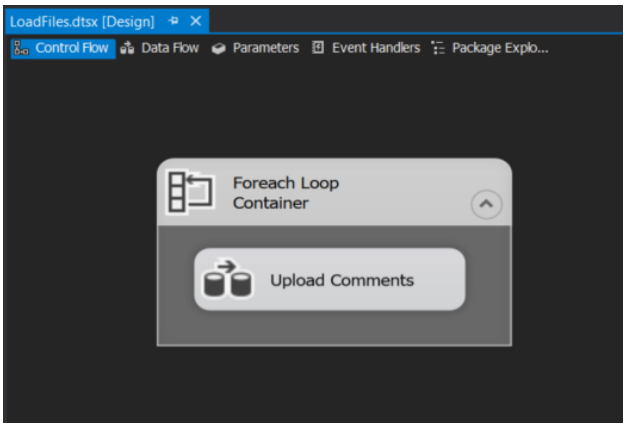


Figura 33 Contenedor de bucle para cada objeto. Fuente: Elaboración propia, 2021.

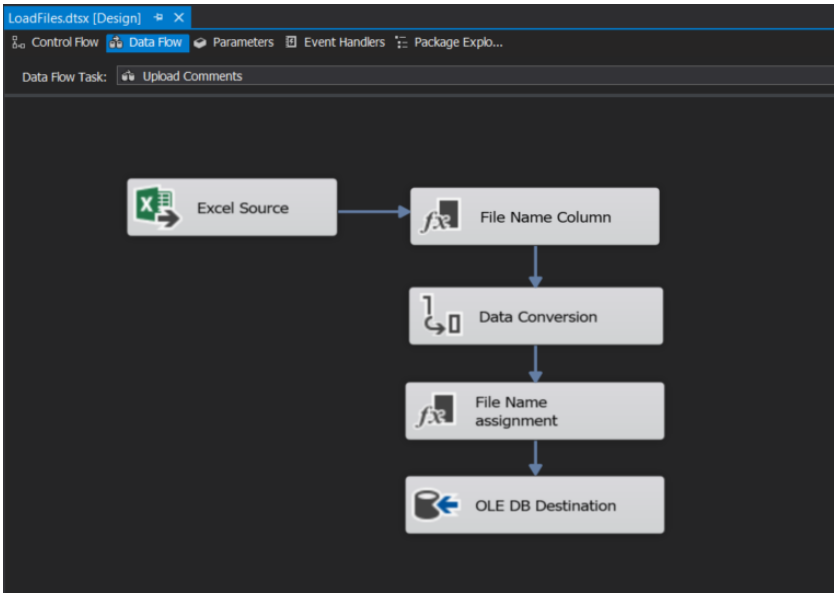


Figura 34 Flujo de datos: transformación y carga a la base de datos. Fuente: Elaboración propia, 2021.

5.1.3.2 Limpieza y transformación.

Para este proyecto, la limpieza consumió la mayoría de tiempo en desarrollarse, puesto que no se puedo utilizar una herramienta de automatización por las características del dataset.

Al contar con varios tipos de recursos/entradas, el trabajo en detención de errores, duplicación e inconsistencias se multiplicó al querer generar un solo conjunto

de datos. La figura 35 sintetiza a través de un diagrama de flujo el proceso de limpieza ejecutado.

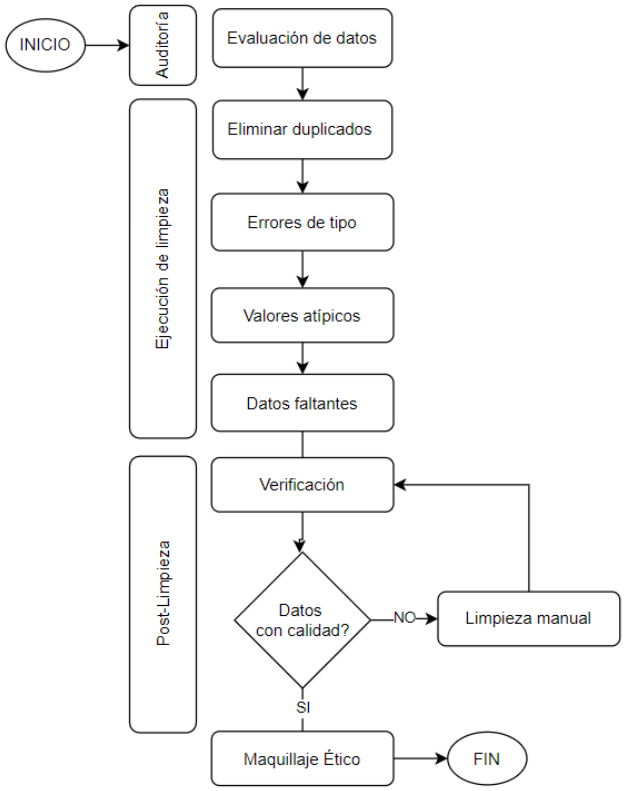


Figura 35 Diagrama de flujo "Limpieza de datos". Fuente: Elaboración propia, 2021.

Se describen de los procesos después de la auditoría de datos a continuación:

5.1.3.3 Eliminar observaciones duplicados.

Al recopilar datos de OCC y OIC por separado se generó una gran cantidad de registros duplicados que debieron ser procesados en tres oportunidades. La primera vez, en la tabla original “ordenes” (mezcla de OCC y OIC) obtenidos directamente del recurso/entrada; la segunda, al separar los registros de la tabla “ordenes” en dos tablas: “OCC” y “OIC”; la tercera y última, combinar los resultados de las tablas OCC y OIC en la tabla final “Prod_analysis”. Vale la pena rescatar que las entregas parciales

en las OIC requirieron un esfuerzo de limpieza manual, la desarrolladora eligió manualmente la observación con mayor detalle de las variables requeridas para este proyecto y eliminaron las observaciones restantes.

5.1.3.4 Errores de tipo.

Nomenclaturas extrañas, utilización de (-) representando No aplica (NA), errores tipográficos, utilización de mayúsculas, exceso de espacios en blanco en cualquier parte de la cadena de letras, números o fechas, diversidad en formatos de fecha, entre otros, fueron las inconsistencias encontradas y corregidas por la desarrolladora.

5.1.3.5 Valores atípicos no deseados.

Un importante hallazgo fue el traslape entre los rangos de fechas de los diversos estatus/ciclos de una sola orden, por ejemplo: una orden reportada con un problema en producción el día 02/16/202, la misma orden el día 01/01/2021 fue reportada con una salida del producto de la fábrica. Para controlar este problema, se decidió tomar los parámetros de producción y aspectos técnicos desde el sistema ERP y la información restante fue tomada directamente de estos reportes generados por el departamento de la cadena de suministro.

La otra sorpresa fue encontrar inconsistencias en el sistema ERP, en menor cantidad, esto requirió la manipulación manual de los datos con ayuda del usuario o ajuste manual según orden cronológico de los eventos. Ver figura 36, la fecha de creación es mayor que la fecha de limpieza.

Sales Doc.	Item	Created on	Clean Date
[redacted]	240	01/18/2021	01/13/2021

Figura 36 Inconsistencia en sistema ERP. Fuente: Sistema ERP

5.1.3.6 Datos faltantes.

Los datos faltantes en las OIC requirieron mayor atención por contener los tiempos de las variables de interés, dos fueron los problemas erradicados en esta tabla.

El primer punto, la pérdida en señal de la orden por días, semanas o quincenas; existen tres conjeturas del porqué, la primera es la pérdida de la orden en el almacén de datos; omisión humana o el cambio del colaborador por tema de vacaciones, incapacidades, etc.; la tabla 23 ejemplifica el hallazgo.

Tabla 23 Comentarios orden A

Orden	Ítem	Comentario	Estatus	Última actualización	Ingresado por
Order A	310	Comentario 200	Producción	03/05/2021	Colaborador 2
Order A	310	Comentario 101	Producción	03/04/2021	Colaborador 2
Order A	310	Comentario 182	Producción	03/03/2021	Colaborador 2
Order A	310	Comentario 225	Producción	03/02/2021	Colaborador 2
Order A	310	Comentario 63	Escasez de material	**03/01/2021**	**Colaborador 1**
Order A	310	Comentario 45	Escasez de material	**02/17/2021**	Colaborador 2
Order A	310	Comentario 150	Escasez de material	02/16/2021	Colaborador 2
Order A	310	Comentario 240	Escasez de material	02/15/2021	Colaborador 2
Order A	310	Comentario 26	Escasez de material	02/11/2021	Colaborador 2
Order A	310	Comentario 165	Escasez de material	02/10/2021	Colaborador 2

Fuente: Reporte de estatus de las órdenes.

El segundo punto corresponde a la pérdida de señal en la salida de producción por ser en fines de semana o una salida extraordinaria del producto y facturación inmediato, ello produce que la orden sea eliminada por la lógica en el almacén de datos. Minimizar este hallazgo representó agregar extracciones adicionales en los sistemas ERP o en el almacén de datos buscando una fecha de salida del OIC, OCC o facturación para completar la información pérdida, se trató de evitar operar desde

suposiciones, pues alteraría las observaciones reales y la pérdida de integridad de datos.

El cruce entre los órdenes OIC y OCC evidenció faltantes de alta criticidad en 100 órdenes, se decidió eliminarlas en ambas tablas y del análisis final, también, los datos de la OIC tenían mayor grado de completitud que la OCC, se procedió con la sobreescritura de herencia sobre los datos perdidos en la OCC.

Para concluir la sesión de datos faltantes, el procesamiento de nulos ameritó la creación de una categoría de agrupación extra solo en el caso de variables cualitativas, previendo fallas en el modelo por concepto de nulos u omisiones de datos.

5.1.3.7 Control de calidad.

Se idearon 5 preguntas para evaluar la calidad de cada una de las variables.

1. ¿Los datos tienen sentido o significancia para este proyecto?
2. ¿Los datos siguen las reglas apropiadas para su campo?
3. ¿Los datos sacan a la luz alguna idea para este proyecto?
4. ¿Se puede generar alguna tendencia a partir de los datos que muestra la variable?

En caso de que la respuesta no fuera clara o negativa, se procedió a realizar una limpieza manual.

El resultado del trabajo en este capítulo se concentró en la tabla 24, los tiempos de producción, escasez de material y redirección exigieron un nivel alto de atención, limpieza y modificación.

Tabla 24 Limpieza por variable expresado en porcentaje

Campo	Campo OIC	Campo OCC
CleaningTime	5%	1%
Redirection	76%	NA
MaterialWT	72%	NA
ProductionTime	**100%**	NA
TestFlag	**100%**	NA
Hold3	1%	NA
Hold6	2%	NA
Hold2	2%	NA
SC Comment Location1	50%	NA

Fuente: Elaboración propia, 2021.

5.1.3.1 Maquillaje ético.

Siguiendo las condiciones de confidencialidad pedidas por el patrocinador, se emplearon técnicas diversas, algunas de ellas fueron descritas en la tabla 25 y las restantes se reservan el derecho de publicación para tranquilidad de nuestro patrocinador. El diseño de las tablas “OIC” y “OCC” contiene la información original sin ninguna transformación; la tabla “*Prod_analysis*” quien consolida los datos del OCC y complementa con detalles del OCC, suprime nombres de clientes, fábricas, configuraciones reales de los productos y los generaliza en grupos según nivel de significancia para este proyecto.

Tabla 25 Técnicas de maquillaje ético

Variable	Técnica
Localization	Se identifican los países con características de entrega complejas y comportamiento similar. Asignación manual de localización. Fórmula: Location + número consecutivo iniciando en 1 según orden de aparición.
DeliveryCondition	Asignación manual de la forma de entrega. Combinación de varios campos del almacén de datos que concluyen la forma de entrega

	de la orden: total, parcial o por número de cotización.
CustomerSegment	Asignación manual por tipo de cliente. Fórmula= CustomerSegment + número consecutivo iniciando en 1 según orden de aparición.
CustomerPriority	Asignación manual de la prioridad de abastecimiento al cliente. Fórmula = CustomerPriority + número consecutivo iniciando en 1 según orden de aparición.
ProductClass	Asignación de la clase del producto manual por grupos. Fórmula= ProductClass + número consecutivo iniciando en 1, orden aleatorio.
Family	Asignación de familia del producto. Fórmula = Family + número consecutivo iniciando en 1, orden aleatorio
ProductType	Asignación según grupo de pertenencia. Fórmula = ProductType + número consecutivo iniciando en 1, orden a conveniencia del proyecto.
Supplier	Asignación manual de fábrica. Fórmula= Supplier + número consecutivo iniciando en 1 según orden de aparición.

Fuente: Elaboración propia, 2021.

5.1.4 Exploración de los datos.

La exploración de datos se dividió en 3 pasos:

1. Exploración de variables
2. Visualización gráfica de datos
3. Análisis de hipótesis del equipo de SC

5.1.4.1 Exploración de variables.

El primer paso, se graficó la relación entre las diferentes variables cuantitativas utilizando la matriz de gráficos de dispersión, ver figura 37.

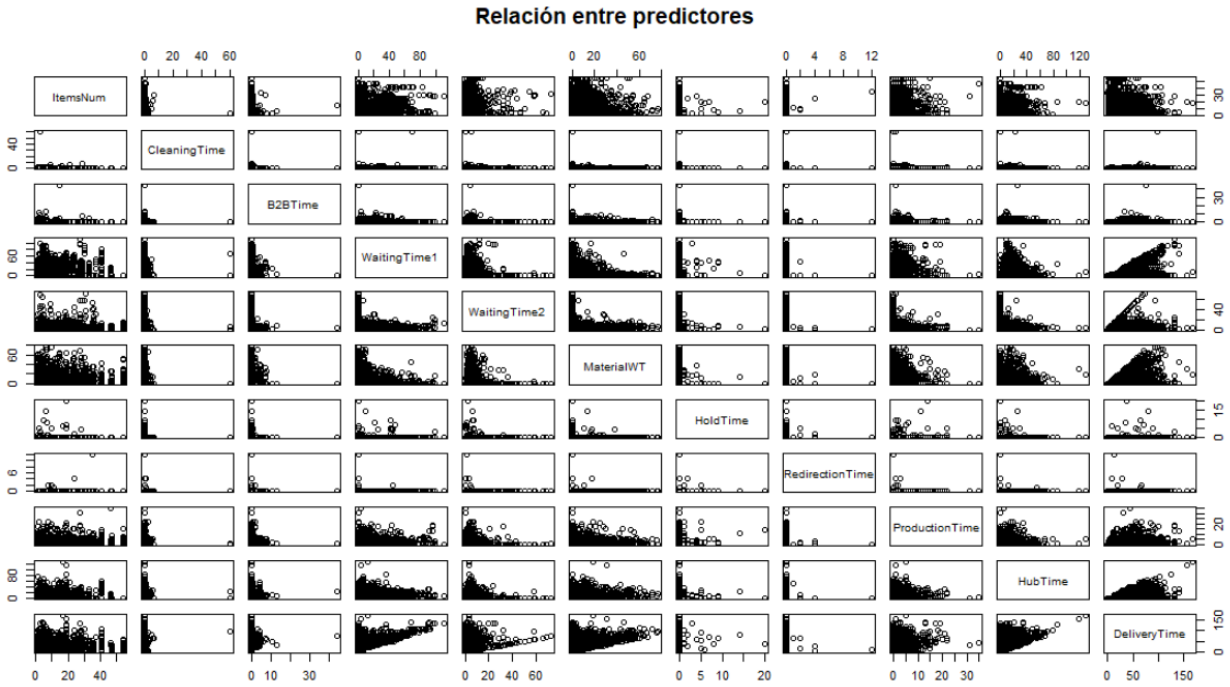


Figura 37 Matriz de gráficos de dispersión. Fuente: Elaboración propia, 2021.

Interpretación: Se observan posibles relaciones positivas fuertes entre las variables *WaitingTime1, WaitingTime2, MaterialWT y HubTime con la variable Delivery Time.* La Figura 38 muestra las leves inclinaciones positivas más valores atípicos a partir de los 100 días en duración de entrega. La presencia de 3 órdenes referentes a la localización 3-1 solicitando productos de alta complejidad en construcción (*Group* 4) y la responsabilidad del Supplier7 cuya entrega al cliente es parcial; son las causas de los valores “anómalos”. Las mencionadas órdenes invierten más de 100 días en *WaitingTime*1 y otros ítems recaen en *HubTime* >100 días, cuyo resultado es la ampliación irregular de los tiempos de entrega.

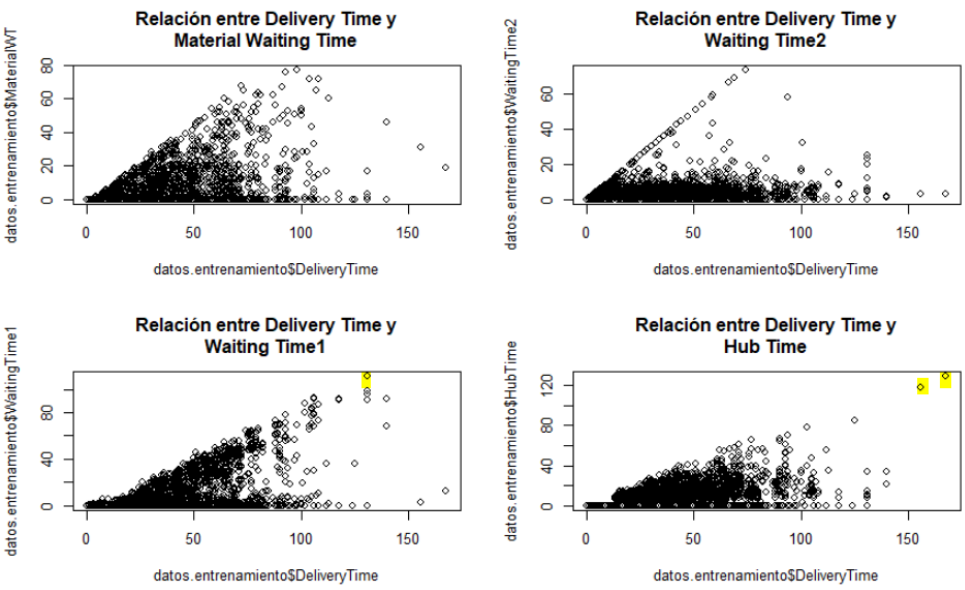


Figura 38 Matriz de dispersión para variables "fuertes". Fuente: Elaboración propia, 2021.

Por su cercanía al valor 0 en el eje “y”, las siguientes variables evidencian debilidad en relación: CleaningTime, B2BTime, RedirectionTime y HoldTime, Fig. 39. Es importante resaltar la presencia de “valores atípicos” falsos, dado que a cada uno de esos puntos corresponde la existencia de un problema en la orden, antes de la etapa de producción. Los puntos extremos resaltados en amarillo pueden ser interpretados con toda seguridad, como la existencia de una excepción.

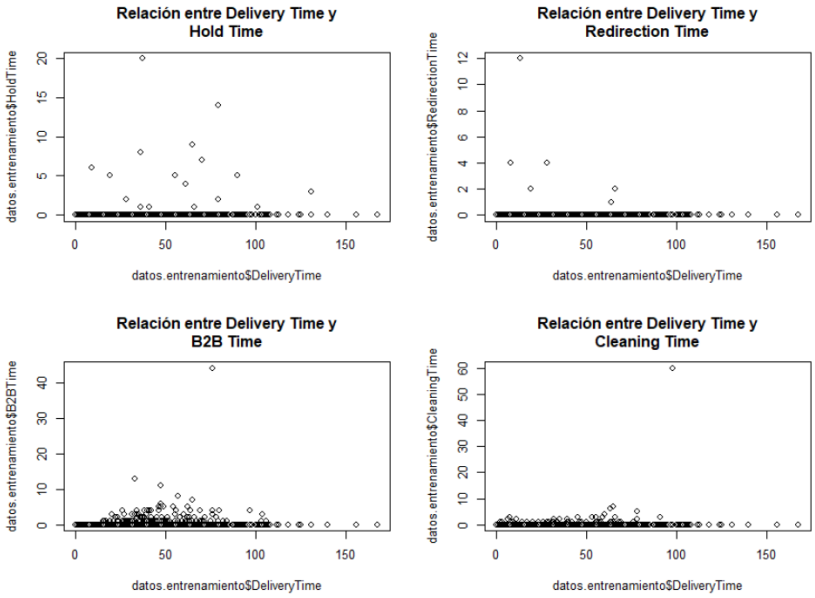


Figura 39 Matriz de dispersión para variables "débiles". Fuente: Elaboración propia.

El siguiente paso fue el estudio de correlación entre todas las variables cuantitativas, la tabla 26 recopila el resultado final; ninguna de las variables es >0.75, la relación entre predictores no muestra señales matemáticas de encontrar un posible patrón.

Tabla 26 Correlación entre variables cuantitativas.

	Items	Cleaning	B2B	Waiting1	Waiting2	Material	Hold	Redirection	Production	Hub	Delivery
ItemsNum	1.00	-0.03	-0.04	-0.05	0.11	-0.06	-0.01	0.04	-0.06	-0.14	-0.14
Cleaning	-0.03	1.00	-0.01	0.04	0.00	0.00	-0.01	0.00	0.00	0.01	0.07
B2B	-0.04	-0.01	1.00	0.07	-0.01	0.04	0.01	0.00	0.03	0.10	0.08
Waiting1	-0.05	0.04	0.07	1.00	-0.03	-0.11	0.06	-0.02	0.12	0.31	0.59
Waiting2	0.11	0.00	-0.01	-0.03	1.00	0.02	0.00	-0.01	-0.04	-0.06	0.04
MaterialWT	-0.06	0.00	0.04	-0.11	0.02	1.00	0.00	-0.01	0.15	0.13	0.24
Hold	-0.01	-0.01	0.01	0.06	0.00	0.00	1.00	0.02	0.15	0.02	0.07
Redirection	0.04	0.00	0.00	-0.02	-0.01	-0.01	0.02	1.00	-0.02	-0.02	-0.03
Production	-0.06	0.00	0.03	0.12	-0.04	0.15	0.15	-0.02	1.00	0.14	0.06
Hub	-0.14	0.01	0.10	0.31	-0.06	0.13	0.02	-0.02	0.14	1.00	0.42
Delivery	-0.14	0.07	0.08	0.59	0.04	0.24	0.07	-0.03	0.06	0.42	1.00

Fuente: Elaboración propia, 2021.

Seguidamente, utilizando la matriz de dispersión se analizan las variables cualitativas, ver figura 40.

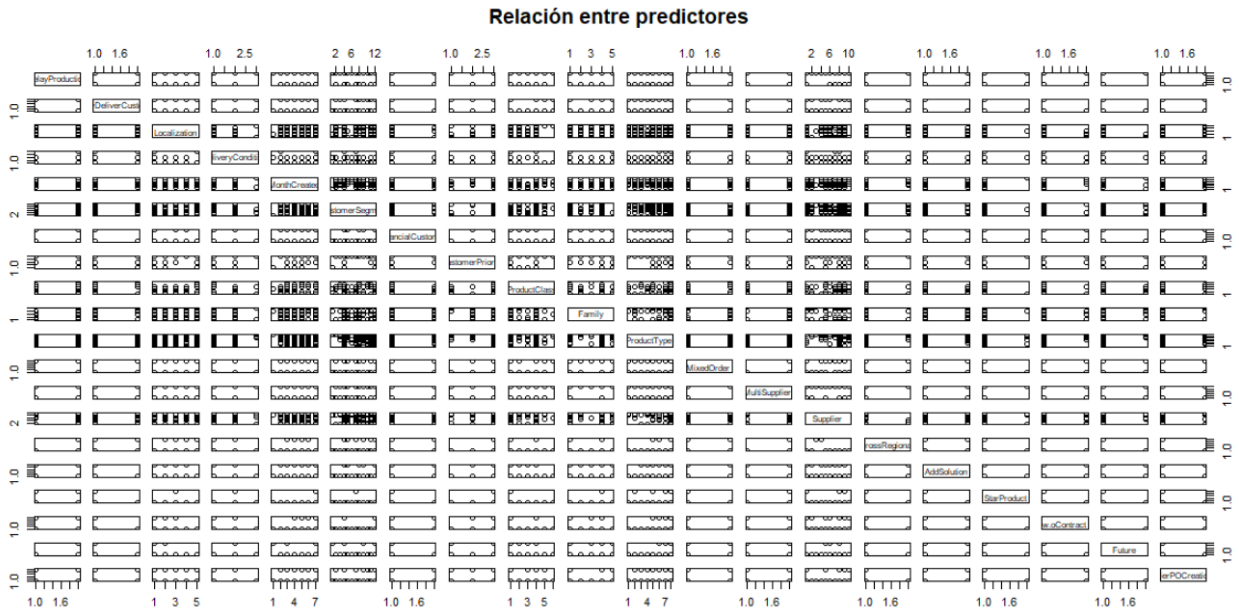


Figura 40 Matriz de gráficos de dispersión. Fuente: Elaboración propia, 2021.

No existen relaciones visibles entre las diferentes variables cualitativas. *MonthCreated, CustomerSegment, ProductType y Supplier* señalan gran dispersión en comparación con el resto que suele agruparse en dos grupos ya sea en el eje “y” o “x”.

Como resultado de esta sesión, se propuso el uso de 15 variables para la construcción de los modelos: *WaitingTime1, WaitingTime2, MaterialWT, HubTime, ProductionTime* y por interés del negocio: *DeliveryCondition, Localization, MultiSupplier, ProductClass, ProductType, Supplier, SupplierPOCreationLate y MixedOrder.*

5.1.4.2 Visualización gráfica de datos.

Los datos recolectados para este proyecto presentan las siguientes características o comportamientos:

En la figura 41, se aprecia el reparto de ítems por mes entre un rango de 1 a 853, los cuadrados naranja indican la cantidad de ítems entregados tarde; solo una orden es entregada a tiempo: mes noviembre, orden perteneciente a la localización 2; para los meses restantes el patrón revela la duplicación o más del número de órdenes tardías en relación con el número de las órdenes a tiempo.



Figura 41 Distribución de ítems con relación a su tiempo de entrega por mes. Fuente: Elaboración propia, 2021.

A nivel general los ítems que no cumplieron la fecha de entrega se extienden a “x” días promedio, mientras que las que cumplieron los parámetros toman tan solo “x” días, según sugiere la figura 42.A.

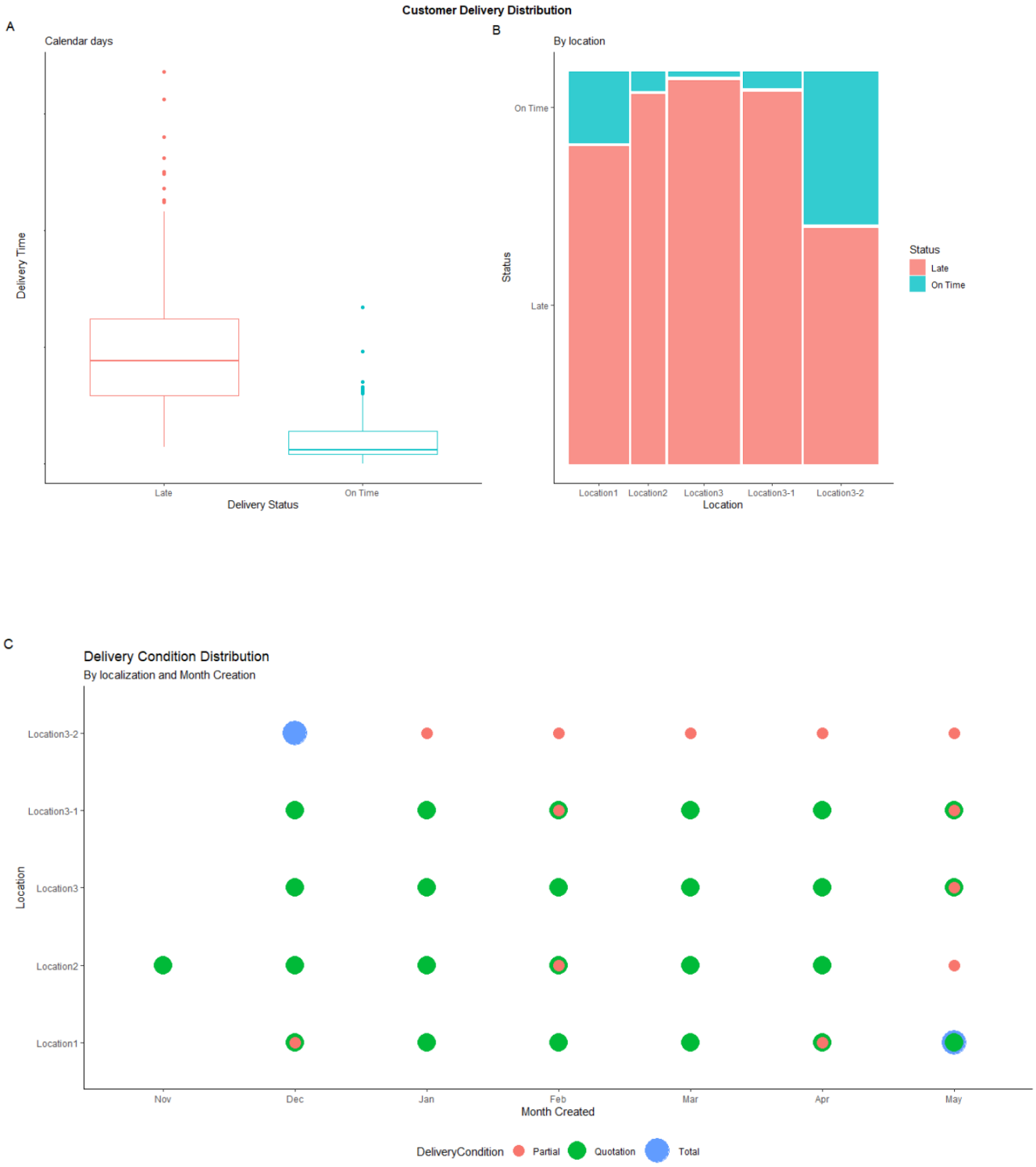
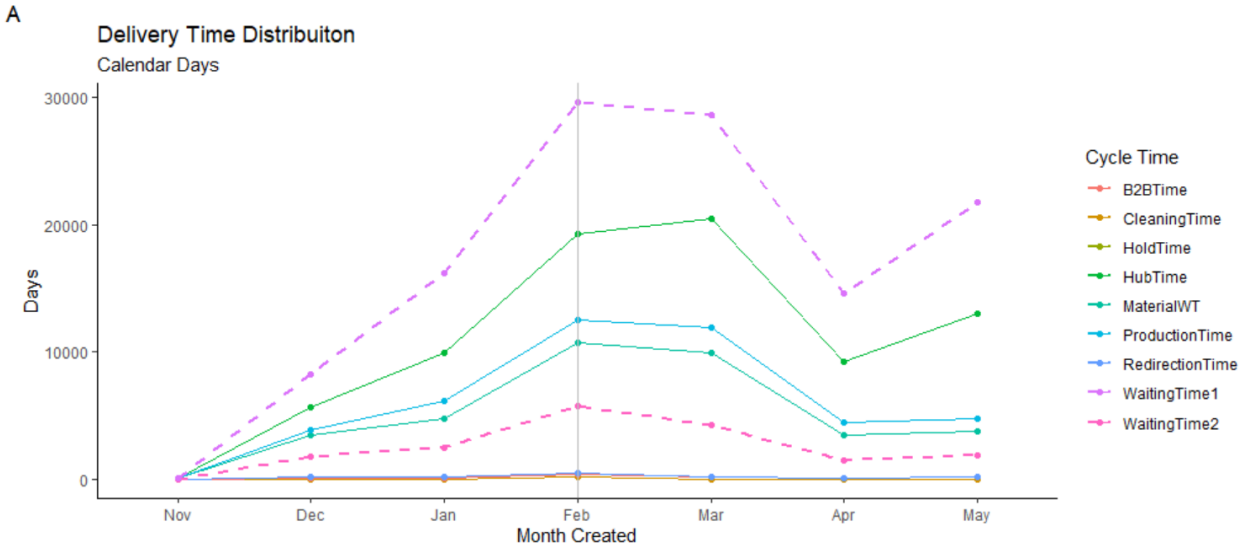


Figura 42 Reparto de los ítems según estado de entrega. Fuente: Elaboración propia, 2022.

A. Promedio de días por estatus de entrega
B. Distribución de ítems por localización y estado de entrega.
C. Distribución de ítems por la condición de entrega por localización.

Según la figura 42.B, en la localización 3-2 predominan entregadas a tiempo gracias a la gran cantidad de ítems en condición de entrega parcial, esta es la misma situación para localización 1. La localización 1 es la única que tiene las tres modalidades de entrega: parcial, total o por agrupamiento del número de cotización. Un 49.14% de las órdenes tienen dependencia al complete de todos los ítems de las órdenes bajo el número de cotización, mientras que el 50.66% están condicionadas a entrega parcial con un restante del 0.19% para órdenes de entrega total.

Por otro lado, queda el estudio de los tiempos, por lo que se crearon dos gráficos más, expuestos en la figura 43, el primer gráfico exhibe la duración devengada de los ítems en diferentes ciclos y el segundo enlista los problemas detectados por mes.



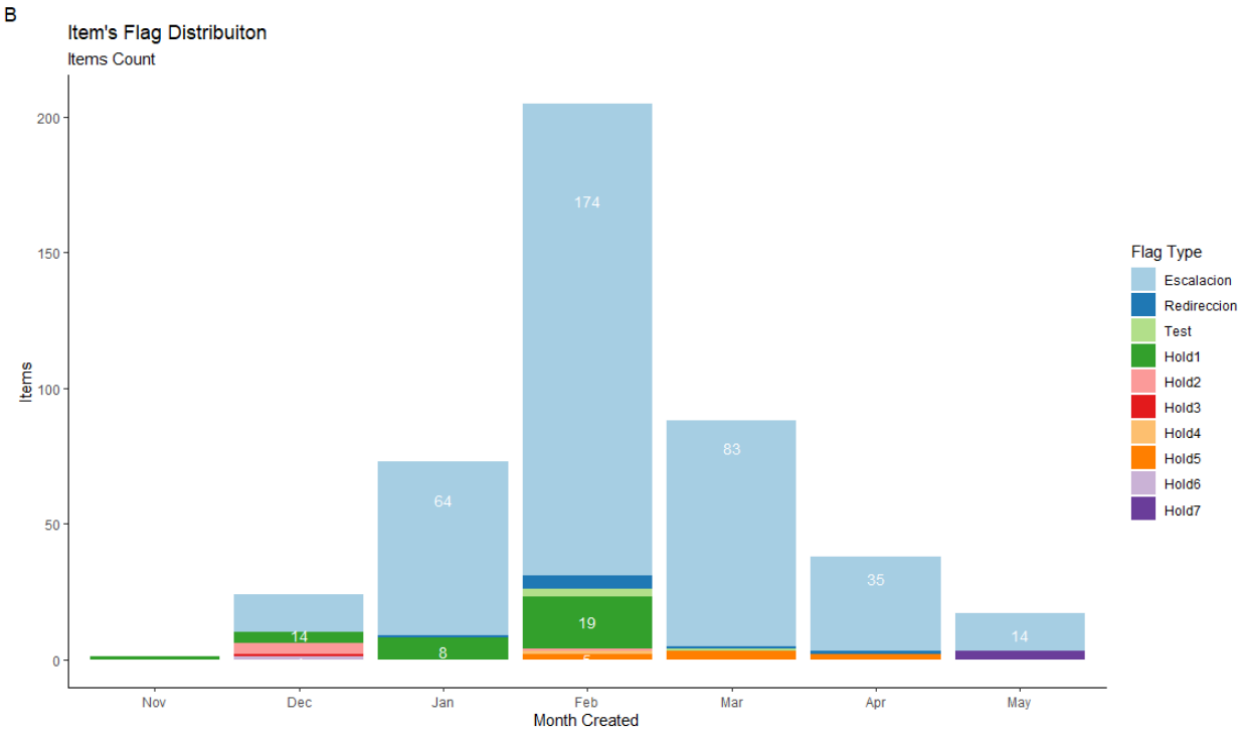


Figura 43 Distribución de los tiempos de entrega según ciclo de vida del ítem. Fuente: Elaboración propia, 2021

En la figura 43.A, la línea gris separa los meses de creación por cuarto, dos son los cuartos utilizados entre el año 2019 (Q4) y 2020 (Q1); los tiempos de las secciones: Material Waiting Time, Waiting Time 2, Hub Time y Waiting Time 1 incrementan exponencialmente en relación con los otros ciclos restantes. Se pueden observar días de espera antes y después de la construcción del producto y la cercanía entre las líneas Material Waiting Time y Producción; es decir, cuando la escasez de material es resuelta es señal casi inmediata de inicio en producción. Las líneas puntiagudas separan los tiempos de espera de los estados normales de la orden, sobresale el Waiting Time 1; ello se atribuye a la duración entre el almacén y la puesta del producto en la puerta del cliente, sin embargo, el criterio experto indica que el número es muy

alto, por lo tanto, se debe hacer un estudio adicional para determinar la anormalidad detectada.

Los meses de febrero y marzo presentan mínimas diferencias mientras que enero y abril expresan patrones parecidos, pero en sentidos opuestos. El declive percibido en el mes de abril fue por el bajo número de órdenes creadas, una situación similar se dio para los meses de noviembre y diciembre.

Febrero vuelve a sobresalir con una presencia fuerte de ítems que fueron escalados (ver figura 43.B). Las posibles causas fueron: no tener respuesta de la fábrica, problemas con recuperación de materiales o fecha de entrega en el pasado; asociando este resultado con la figura 43.A, la sección de escasez en materiales incrementó en relación con el mes de enero. Seguidamente el gráfico de barras sugiere problemas con el hold1, redirección y hold5.

5.1.4.3 Análisis de hipótesis del equipo de SC.

A continuación, se recapitulan y analizan las teorías usando los datos recolectados con su respectivo gráfico según tipo de dato.

1. Capacidad de fábrica no es el problema

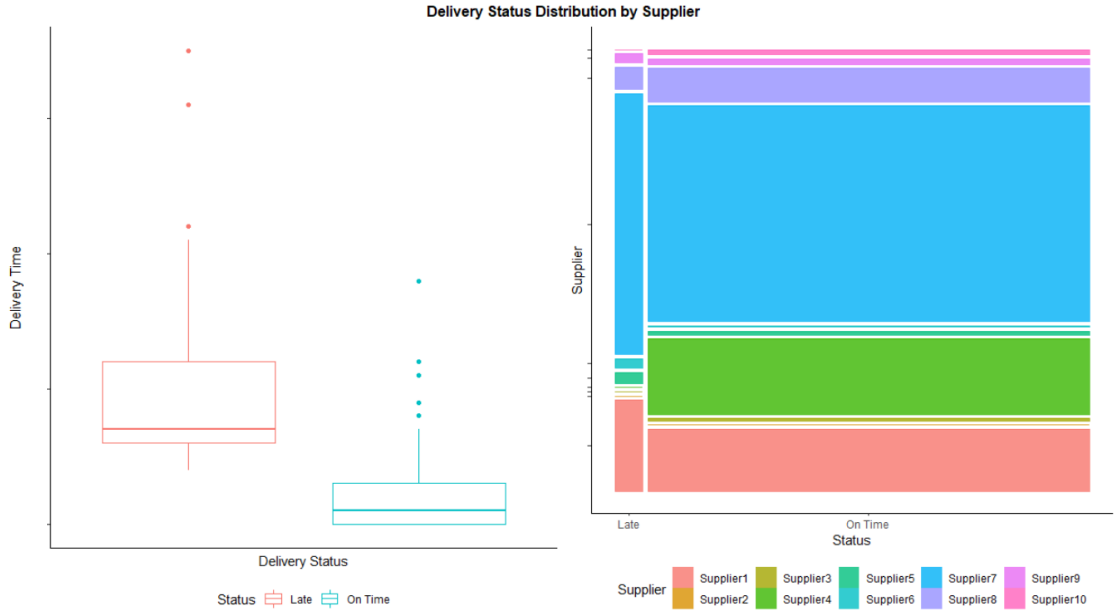


Figura 44 Distribución del tiempo de entrega de fábrica. Fuente: Elaboración propia, 2021

Resultado fig. 44: Verdadero, tanto los tiempos de entrega a tiempo como tardío son < “x” días de entrega. Fábrica 7 tiene el mejor récord de cumplimiento contra los tiempos estándar.

Alrededor de un 33% de los casos tardíos fueron escalados, 22% presentaron problemas en *testing* y 38% hold1, pero no se observa una causa en común entre todos los ítems. En cuanto a los puntos fuera del promedio de los ítems a tiempo, 100% están atados a entrega por cotización, 73% con hold 1 y un 36% fue escalado.

2. Productos con solución especial siempre se entregan tarde.

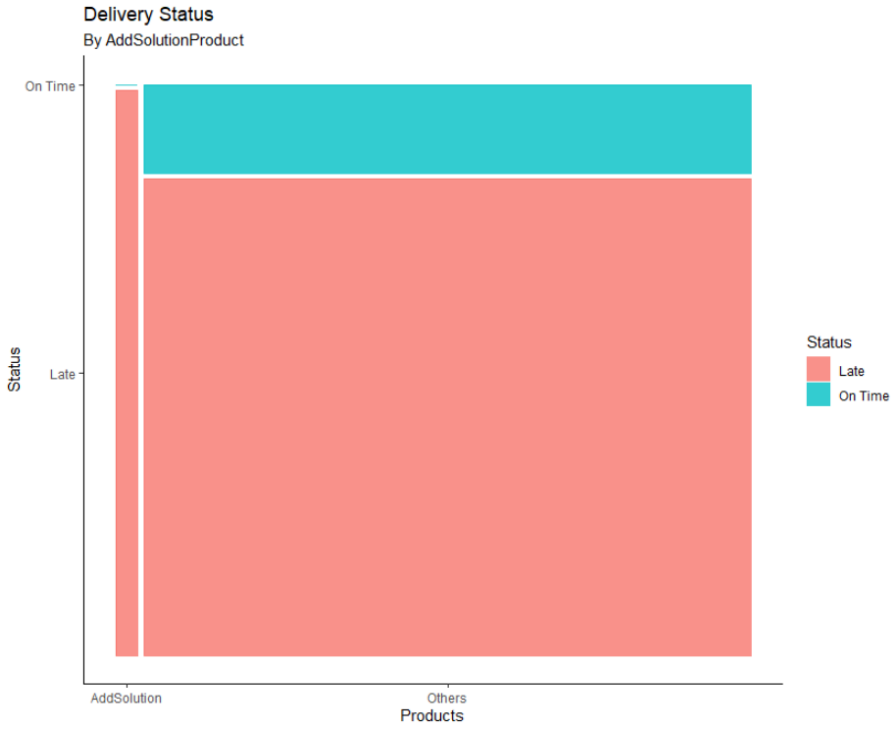


Figura 45 Entrega de producto con soluciones adicionales. Fuente: Elaboración propia, 2021

Resultado fig. 45: Verdadero, los productos que requieren una solución adicional son dados tarde al cliente el 98% de los casos.

3. Escasez en productos de gran valor (Grupo 1, 2 y 3) son causantes de tardías

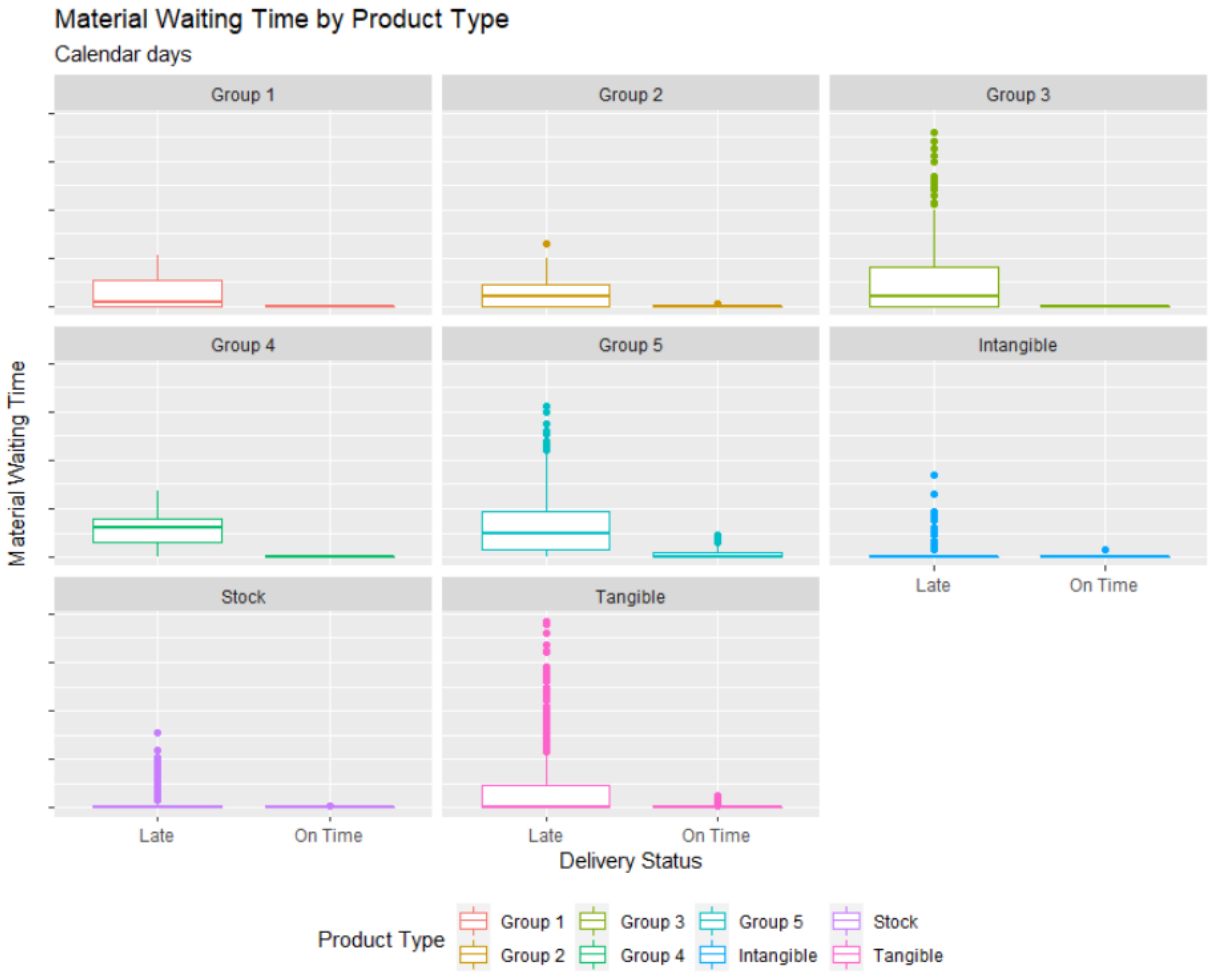


Figura 46 Tiempo de recuperación de materiales por tipo de producto. Fuente: Elaboración propia, 2021.

Resultado fig. 46: Parcialmente verdadero. Analizando el promedio del tiempo en recuperación de los *product type*, los grupos 3 y 5 presentan mayor cantidad de días; no se puede dejar de mencionar que el total en promedio no excede a los “x” días exceptuando algunos casos. El comportamiento de los productos tangibles está regido por una gran diversidad de productos que explica el puntillismo fuera del promedio en la sesión tarde del gráfico. Seguidamente, el puntillismo del grupo 3 se asemeja por el uso de fábricas no oficiales y son justamente los productos con cortos a nivel mundial.

4. Estados de la orden bloqueo 5, 6 y 7 aumentan el tiempo de entrega

Los datos colectados no tienen un porcentaje de significancia en el total del tiempo de entrega (0.15%), razón la cual, no se profundiza en este estudio.

5. Productos estrella mejoran los tiempos de entrega

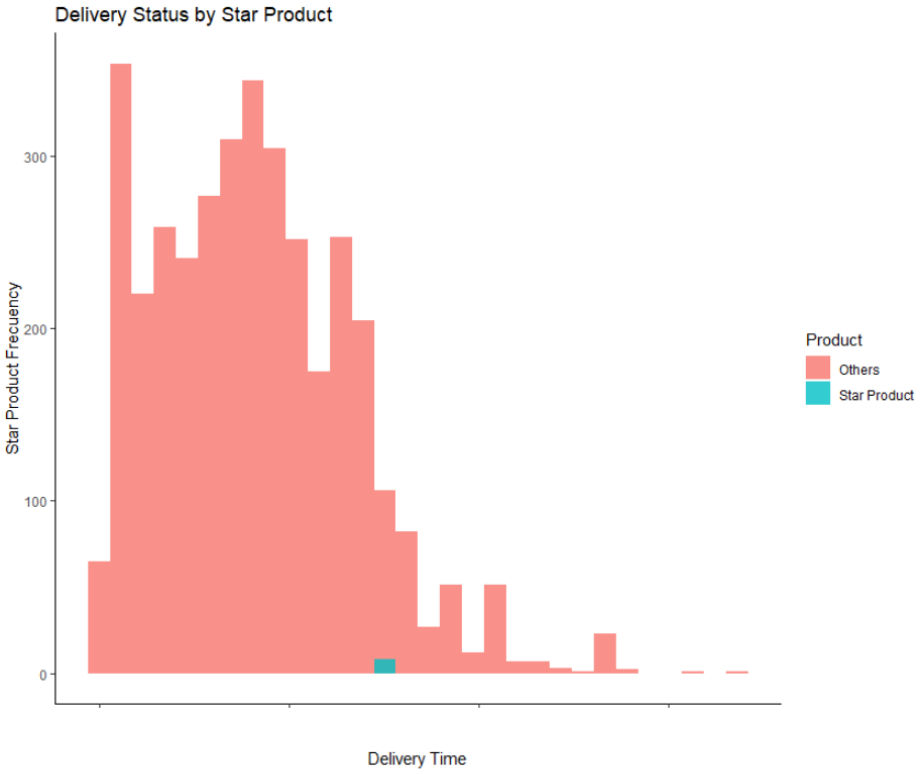


Figura 47 Estatus de entrega por productos estrella. Fuente: Elaboración propia, 2021.

Resultado fig. 47: El *set* de datos cuenta con una orden con productos estrella, no es suficiente para declarar una hipótesis verdadera o falsa, ahora, el análisis dio un total de “x” días en entrega, este sería el punto de inicio para investigaciones futuras.

6. El problema es el tiempo de consolidación de productos

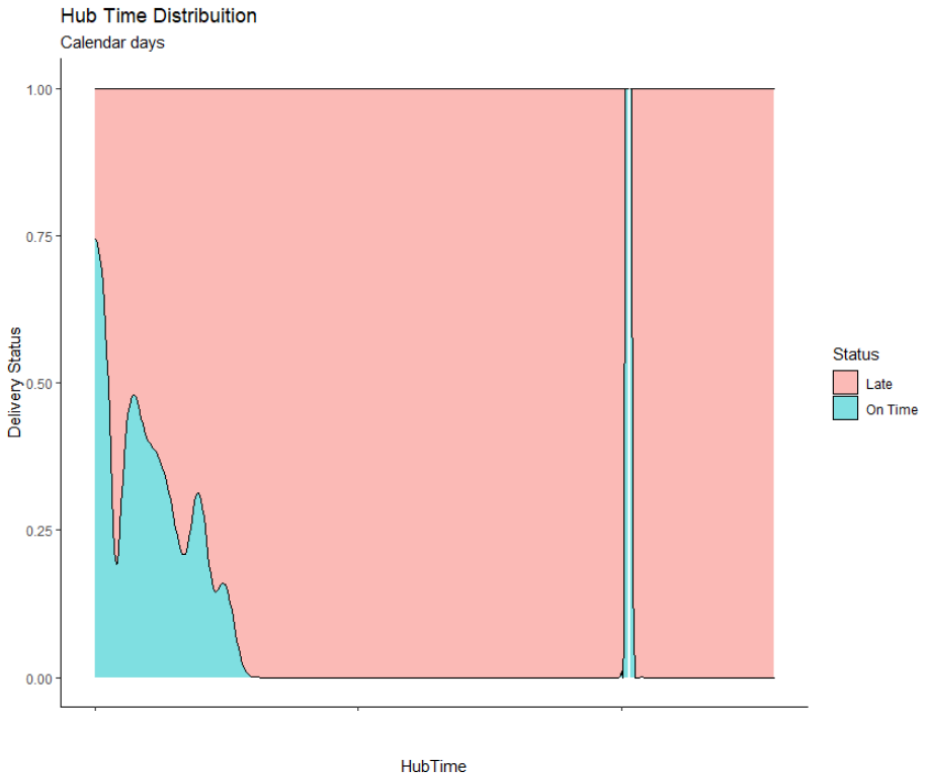


Figura 48 Distribución de los tiempos en Hub. Fuente: Elaboración propia, 2021.

Resultado fig. 48: Verdadera. La zona naranja afirma que el tiempo de consolidación es una de las causales en incumplimiento de entrega.

7. Aumentar la prioridad de la orden (nivel 3), aumenta la posibilidad de entrega a tiempo

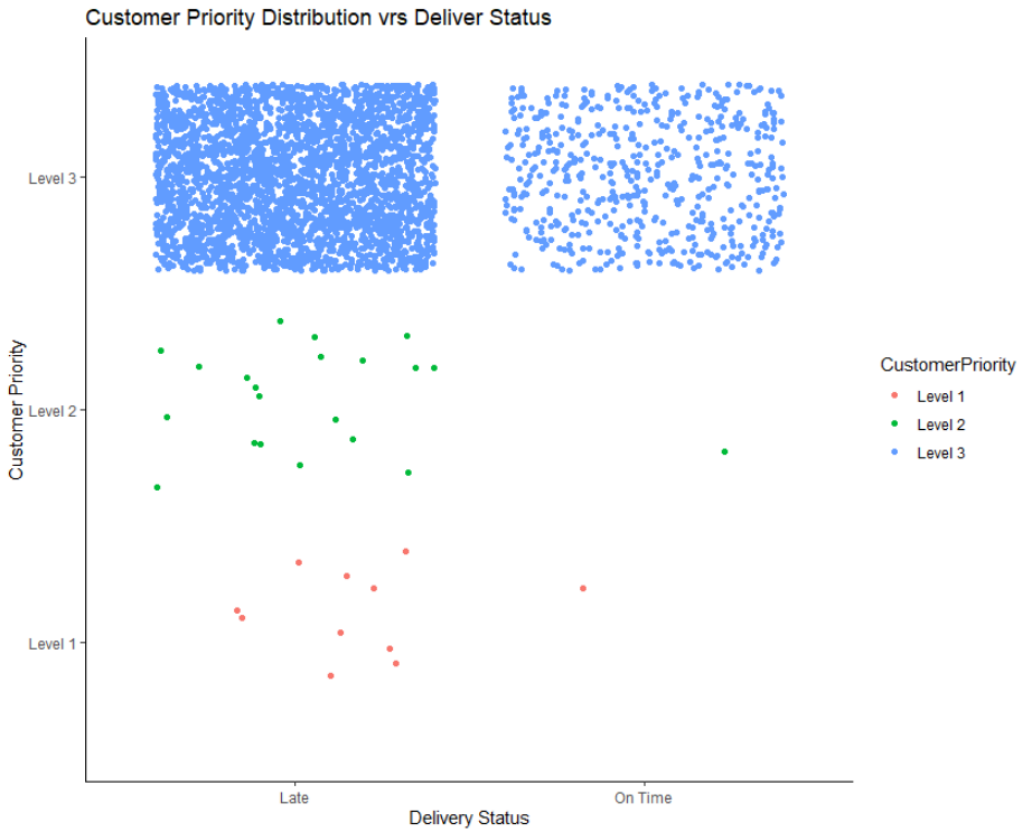


Figura 49 Distribución de entrega según la prioridad del cliente. Fuente: Elaboración propia, 2021.

Resultado fig. 49: Esta teoría no siempre es verdadera. Estudiando el cuadrante superior derecho, solo el 18% del total de las órdenes en nivel 3 lograron mejorar la posibilidad de cumplimiento.

8. Problemas con los compradores internos

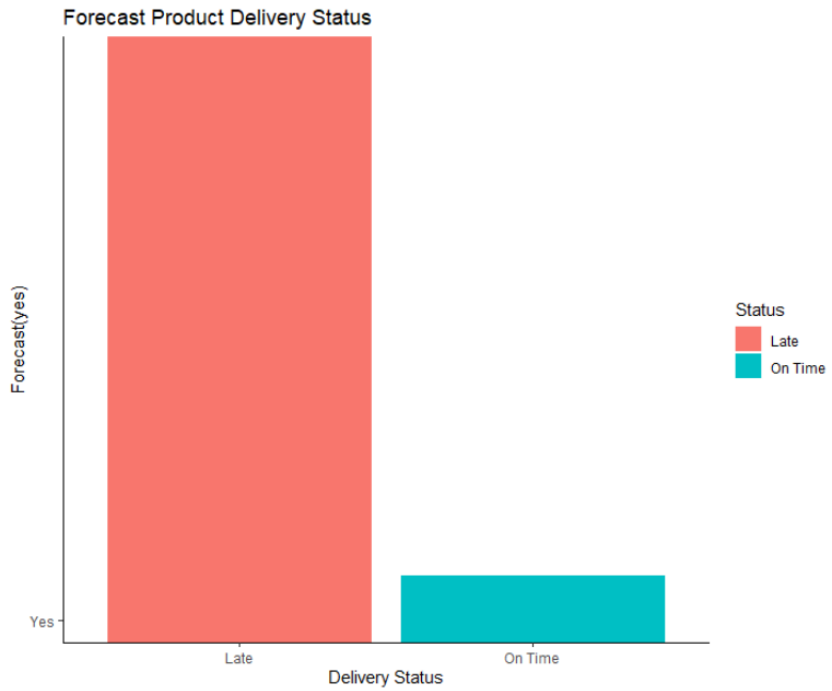


Figura 50 Distribución del estatus de entrega en productos incluidos en la planeación.
Fuente: Elaboración propia, 2021.

Resultado fig. 50: Falso. Los datos extraídos remarcan solo un 9.76% de efectividad. Lastimosamente, el dataset solo cuenta con datos en la localización 1.

5.1.5 Modelado.

R Studio es una poderosa herramienta estadística que permitió realizar con facilidad los análisis requeridos en este capítulo.

Los experimentos usaron las variables definidas en la sesión 5.1.4.1 y se fueron eliminando según su grado de significancia en el modelo. La tabla 27 resume las variables escogidas y clasificadas por modelos (columnas). Los datos fueron divididos en dos partes: 70% para entrenamiento y 30% pruebas de predicción. El comando

"*train*" facilitó aplicar los mismos métodos y parámetros de validación cruzada y muestreo.

El algoritmo regresión logística y árboles de decisión fueron seleccionados por su alta popularidad y acogimiento por parte del negocio. Adicionalmente, la regresión logística permite ver el comportamiento de variables cualitativas, mientras que la ventaja de los árboles de decisión es la visualización intuitiva del flujo de los resultados, lo cual facilita el entendimiento por parte del patrocinador.

Tabla 27 Variables por modelos

Variable	Modelo 1	Modelo 2	Modelo 3
B2BTime	X	X	X
DeliveryCondition	X		X
HubTime	X	X	
Localization	X		X
MaterialWT	X	X	X
MultiSupplier	X	X	X
ProductClass	X		X
ProductionTime	X	X	X
ProductType	X	X	
Supplier	X	X	X
SupplierPOCreationLate	X		X
WaitingTime1	X	X	
WaitingTime2	X	X	X
MixedOrder			X
DelayDeliverCustomer	X	X	X

Fuente: Elaboración propia, 2021.

Seguidamente se detalla el uso de las dos técnicas de minería:

5.1.5.1 Regresión logística.

El entrenamiento de los 3 modelos generó la figura 51, en donde la métrica valor estimado de la información "perdida": AIC acrónimo en inglés, es de: 981.47, 1148.2 y 1310.3 respectivamente.

Es importante mencionar cómo existencia del quinteto B2BTime, WaitingTime1, WaitingTime2, MaterialWT, ProductionTime y HubTime sobresale en el modelo y disminuye considerablemente el valor de AIC.

Modelo 1

```
Coefficients:
                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                10.57973    1.03642  10.208  < 2e-16 ***
LocalizationLocation2      -2.94292    0.39637  -7.425 1.13e-13 ***
LocalizationLocation3      -1.99967    0.45368  -4.408 1.04e-05 ***
LocalizationLocation3-1     1.91157    0.47181   4.052 5.09e-05 ***
LocalizationLocation3-2    -3.03302    0.54971  -5.518 3.44e-08 ***
DeliveryConditionQuotation -0.58353    0.31054  -1.879 0.060229 .
DeliveryConditionTotal      1.73288    3.27057   0.530 0.596222
ProductClassProductClass2 -13.48592  747.97223  -0.018 0.985615
ProductClassProductClass3  -6.66053  636.84981  -0.010 0.991655
ProductClassProductClass4  -1.87697    0.25400  -7.390 1.47e-13 ***
ProductClassProductClass5  -3.19099    0.62003  -5.146 2.65e-07 ***
ProductClassProductClass6  -2.47903    2.28780  -1.084 0.278547
ProductTypeGroup 2          0.59507    0.94471   0.630 0.528763
ProductTypeGroup 3         -0.12960    0.93132  -0.139 0.889324
ProductTypeGroup 4          3.45587    6.54103   0.528 0.597265
ProductTypeGroup 5          2.69281    0.81661   3.298 0.000975 ***
ProductTypeIntangible      -1.00342    0.89926  -1.116 0.264498
ProductTypeStock            0.12997    0.73197   0.178 0.859066
ProductTypeTangible        -0.07815    0.68809  -0.114 0.909578
MultiSupplierYes           -0.65402    0.26037  -2.512 0.012008 *
SupplierSupplier10         -9.32389    0.79119 -11.785  < 2e-16 ***
SupplierSupplier2          -9.24004 6522.63865  -0.001 0.998870
SupplierSupplier3           2.38627    1.36813   1.744 0.081127 .
SupplierSupplier4         -10.52381    1.08620  -9.689  < 2e-16 ***
SupplierSupplier5           0.95500    1.81586   0.526 0.598944
SupplierSupplier6          -8.96733 1009.35275  -0.009 0.992912
SupplierSupplier7           0.27349    0.26130   1.047 0.295245
SupplierSupplier8           0.78036    0.42037   1.856 0.063404 .
SupplierSupplier9          -0.71512    1.17993  -0.606 0.544467
SupplierPOCreationLateYes  -1.87479    1.07073  -1.751 0.079956 .
B2BTime                    -1.14352    0.32070  -3.566 0.000363 ***
WaitingTime1               -0.23542    0.02601  -9.053  < 2e-16 ***
WaitingTime2               -0.75068    0.06055 -12.397  < 2e-16 ***
MaterialWT                 -0.81293    0.07542 -10.778  < 2e-16 ***
ProductionTime             -0.65425    0.06503 -10.061  < 2e-16 ***
HubTime                    -0.40718    0.04122  -9.878  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3102.16  on 3641  degrees of freedom
Residual deviance:  909.47  on 3606  degrees of freedom
AIC: 981.47

Number of Fisher Scoring iterations: 17
```

Modelo 2

```
Coefficients:
                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                 4.69664    0.69598   6.748 1.50e-11 ***
MultiSupplierYes           -1.02924    0.23317  -4.414 1.01e-05 ***
ProductTypeGroup 2          0.56165    0.80448   0.698 0.485080
ProductTypeGroup 3          0.17228    0.83369   0.207 0.836284
ProductTypeGroup 4          2.07400    2.92329   0.709 0.478030
ProductTypeGroup 5          2.20866    0.65114   3.392 0.000694 ***
ProductTypeIntangible      -1.06687    0.76683  -1.391 0.164142
ProductTypeStock            0.76732    0.62899   1.220 0.222496
ProductTypeTangible         0.40414    0.58657   0.689 0.490828
SupplierSupplier10         -5.48495    0.53649 -10.224  < 2e-16 ***
SupplierSupplier2          -9.61112 2399.54480  -0.004 0.996804
SupplierSupplier3           4.33905    1.14189   3.800 0.000145 ***
SupplierSupplier4          -6.20122    0.63192  -9.813  < 2e-16 ***
SupplierSupplier5           0.26399    0.83280   0.317 0.751251
SupplierSupplier6          -5.60756  372.56030  -0.015 0.987991
SupplierSupplier7           0.59072    0.23243   2.541 0.011039 *
SupplierSupplier8           1.14332    0.38548   2.966 0.003018 **
SupplierSupplier9           2.99381    0.96333   3.108 0.001885 **
B2BTime                    -0.79533    0.25611  -3.105 0.001900 **
WaitingTime1               -0.22032    0.02126 -10.361  < 2e-16 ***
WaitingTime2               -0.66295    0.05314 -12.477  < 2e-16 ***
MaterialWT                 -0.73221    0.07120 -10.284  < 2e-16 ***
ProductionTime             -0.52728    0.05654  -9.325  < 2e-16 ***
HubTime                    -0.15460    0.01535 -10.072  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3102.2  on 3641  degrees of freedom
Residual deviance: 1100.2  on 3618  degrees of freedom
AIC: 1148.2

Number of Fisher Scoring iterations: 15
```

Modelo 3

```
Coefficients:
                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                 3.55167    0.35932   9.884  < 2e-16 ***
LocalizationLocation2      -1.85978    0.30043  -6.190 6.00e-10 ***
LocalizationLocation3      -2.87702    0.35338  -8.142 3.90e-16 ***
LocalizationLocation3-1    -3.12742    0.28448 -10.994  < 2e-16 ***
LocalizationLocation3-2     2.13062    0.24675   8.635  < 2e-16 ***
DeliveryConditionQuotation -1.43453    0.23397  -6.131 8.72e-10 ***
DeliveryConditionTotal      0.20896    1.47672   0.142 0.887470
ProductClassProductClass2 -14.95579  861.02866  -0.017 0.986142
ProductClassProductClass3 -12.55314  671.99137  -0.019 0.985096
ProductClassProductClass4  -1.45278    0.20254  -7.173 7.34e-13 ***
ProductClassProductClass5  -1.08361    0.48314  -2.243 0.024907 *
ProductClassProductClass6  -1.22180    1.68588  -0.725 0.468619
MixedOrderYes              -0.85519    0.43054  -1.986 0.046997 *
MultiSupplierYes           -0.64249    0.21074  -3.049 0.002299 **
SupplierSupplier10         -1.99220    0.41267  -4.828 1.38e-06 ***
SupplierSupplier2         -12.54280 6522.63862  -0.002 0.998466
SupplierSupplier3           2.33597    1.15241   2.027 0.042658 *
SupplierSupplier4          -2.50962    0.36666  -6.845 7.67e-12 ***
SupplierSupplier5           1.10359    1.19748   0.922 0.356741
SupplierSupplier6          -9.85236 1098.11875  -0.009 0.992841
SupplierSupplier7           0.33098    0.20619   1.605 0.108444
SupplierSupplier8           1.40658    0.30674   4.586 4.53e-06 ***
SupplierSupplier9           2.79404    0.82596   3.383 0.000718 ***
SupplierPOCreationLateYes  -1.78058    0.89317  -1.994 0.046202 *
B2BTime                    -1.09417    0.26039  -4.202 2.64e-05 ***
WaitingTime2               -0.53975    0.04446 -12.139  < 2e-16 ***
MaterialWT                 -0.60035    0.06289  -9.545  < 2e-16 ***
ProductionTime             -0.45470    0.04969  -9.151  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

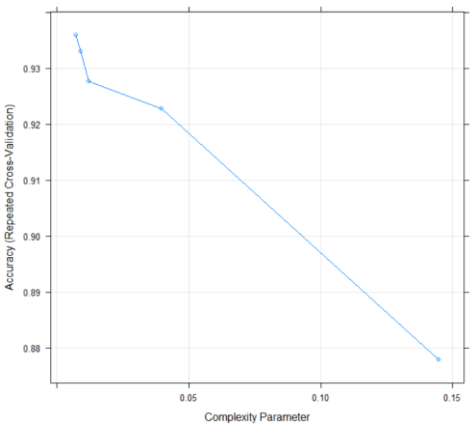
    Null deviance: 3102.2  on 3641  degrees of freedom
Residual deviance: 1254.3  on 3614  degrees of freedom
AIC: 1310.3
```

Figura 51 Resumen de los modelos utilizados en regresión logística. Fuente: Elaboración propia, 2021

5.1.5.2 Árbol de decisión.

Utilizando los mismos modelos, pero cambiando su método de entrenamiento por “rpart”, en figura 52, se aprecia la escogencia del mejor parámetro, coincide en una precisión de un 94% en los tres modelos. El gráfico del lado izquierdo muestra el comportamiento de precisión durante la validación cruzada.

Modelo 1




```
CART

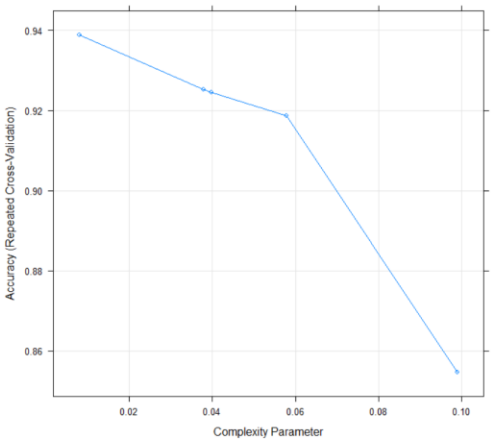
3642 samples
  13 predictor
   2 classes: 'Late', 'On Time'

No pre-processing
Resampling: Cross-Validated (5 fold, repeated 2 times)
Summary of sample sizes: 2914, 2914, 2913, 2913, 2914, 2913,
Resampling results across tuning parameters:

  cp      Accuracy  Kappa
  0.0072  0.94      0.73
  0.0090  0.93      0.70
  0.0121  0.93      0.67
  0.0398  0.92      0.65
  0.1447  0.88      0.26

Accuracy was used to select the optimal model using the larg
The final value used for the model was cp = 0.0072.
```

Modelo 2




```
CART

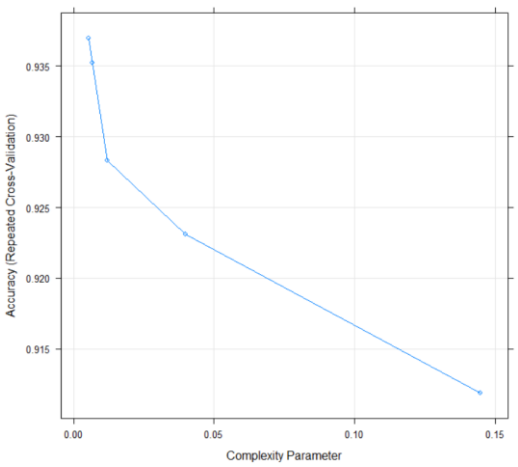
3642 samples
   9 predictor
   2 classes: 'Late', 'On Time'

No pre-processing
Resampling: Cross-Validated (5 fold, repeated 2 times)
Summary of sample sizes: 2915, 2913, 2914, 2913, 2913, 2914,
Resampling results across tuning parameters:

  cp      Accuracy  Kappa
  0.0081  0.94      0.749
  0.0380  0.93      0.692
  0.0398  0.92      0.691
  0.0579  0.92      0.673
  0.0989  0.85      0.064

Accuracy was used to select the optimal model using the larg
The final value used for the model was cp = 0.0081.
```

Modelo 3




```
CART

3642 samples
  11 predictor
   2 classes: 'Late', 'On Time'

No pre-processing
Resampling: Cross-Validated (5 fold, repeated 2 times)
Summary of sample sizes: 2914, 2913, 2914, 2914, 2913, 2914,
Resampling results across tuning parameters:

  cp      Accuracy  Kappa
  0.0054  0.94      0.72
  0.0068  0.94      0.71
  0.0121  0.93      0.67
  0.0398  0.92      0.65
  0.1447  0.91      0.57

Accuracy was used to select the optimal model using the larg
The final value used for the model was cp = 0.0054.
```

Figura 52 Resultado del árbol de decisión en los 3 modelos. Fuente: Elaboración propia, 2021

5.1.6 Evaluación.

Las métricas de selección empleadas fueron sensibilidad (verdaderos positivos), especificidad (verdaderos negativos) y exactitud; utilizando la rúbrica definida en capítulo #3, se tabulan los resultados en la tabla 28.

Tabla 28 Resultados evaluación de modelo y algoritmo.

Método	Modelo	Precisión	Sensibilidad	Especificidad
Regresión Logística	1	95.82	97.65	85.59
	2	94.79	97.05	82.20
	3	93.25	96.29	76.27
Árbol de decisión	1	93.57	98.56	65.68
	2	94.09	96.74	79.24
	3	93.89	97.88	71.61

Fuente: Elaboración propia, 2021

Los mejores resultados lo dio el método regresión logística, modelo #1, con una calificación = Excelente por rendir mediciones >75. En el caso del árbol de decisión el modelo #2 sigue estando dentro de la categoría excelente con el detalle de un evidente bajonazo en la especificidad.

5.2 Propuesta de alteración a la metodología del cálculo de tiempos de entrega

En esta segunda entrega de la solución y aprovechando el arduo trabajo en recolección de datos y análisis de la primera etapa, se procedió a estimar el promedio en días calendario de: MaterialWT, ProductionTime, HubTime, WaitingTime1 y WaitingTime 2, en la tabla 29.

Tabla 29 Resumen estadístico de factores principales de atraso

	MaterialWT	ProductionTime	HubTime	WaitingTime1	WaitingTime2
Promedio	X	X	X	X	X
Desviación estándar	X	X	X	X	X
Mínimo	X	X	X	X	X
Máximo	X	X	X	X	X

Fuente: Elaboración propia, 2021.

Nota: Promedios suprimidos del documento por solicitud del patrocinador.

Sin embargo, cada una de estas etapas presentan variantes condicionales a su cálculo, ver tabla 30.

Tabla 30 Variantes de los tiempos estándar.

Material WT	Avg	Min	Max
ProductClass1	X	X	X
ProductClass2	X	X	X
ProductClass3	X	X	X
ProductClass4	X	X	X
ProductClass5	X	X	X
Total			

Production	Avg	Min	Max
ProductClass1	X	X	X
ProductClass2	X	X	X
ProductClass3	X	X	X
ProductClass4	X	X	X
ProductClass5	X	X	X
Total			

HubTime	Avg	Min	Max
Partial			
Location1	X	X	X
Location2	X	X	X
Location3	X	X	X
Location3-1	X	X	X
Quotation			
Location1	X	X	X
Location2	X	X	X
Location3	X	X	X
Total			
Location1	X	X	X
Location3	X	X	X
Total			

WaitingTime1	Avg	Min	Max
Partial			
Location1	X	X	X
Location2	X	X	X
Location3	X	X	X
Location3-1	X	X	X
Quotation			
Location1	X	X	X
Location2	X	X	X
Location3	X	X	X
Total			
Location1	X	X	X
Location3	X	X	X
Total			

Fuente: Elaboración propia, 2021.

Nota: Promedios suprimidos del documento por solicitud del patrocinador.

Por otro lado, a pesar de la existencia de tantas variables que pudieron justificar la causa de una entrega tardía, la existencia de estos estatus en el set de datos no excede al 14%; es decir, no alcanzó el nivel de importancia en este proyecto; por lo demás, los días en resolución de estos estatus es de “x” o menos en promedio.

En la tabla 31, se resumen dichas variables estadísticamente para uso del negocio, cada vez que una orden presente alguno de estos estatus, se recomienda adicionar estos tiempos a su cálculo estándar.

Tabla 31 Resumen estadístico de variables extra.

	B2B	Hold1	Hold2	Hold3	Hold4	Hold6	Hold7	Redirection	Test fail
Promedio	X	X	X	X	X	X	X	X	X
Desviación estándar	X	X	X	X	X	X	X	X	X
Mínimo	X	X	X	X	X	X	X	X	X
Máximo	X	X	X	X	X	X	X	X	X

Fuente: Elaboración propia, 2021.

Nota: Promedios suprimidos del documento por solicitud del patrocinador.

La metodología del cálculo de tiempo de entrega inicial estaba incompleta, al descartar etapas del ciclo de producción solo porque causaban ruido y enfatizaban las deficiencias del accionar del negocio. Retomando el mapa de flujo de valor expuesto en la sesión 4.3 más los hallazgos de este proyecto, se propone la siguiente reorganización ver fig. 53.

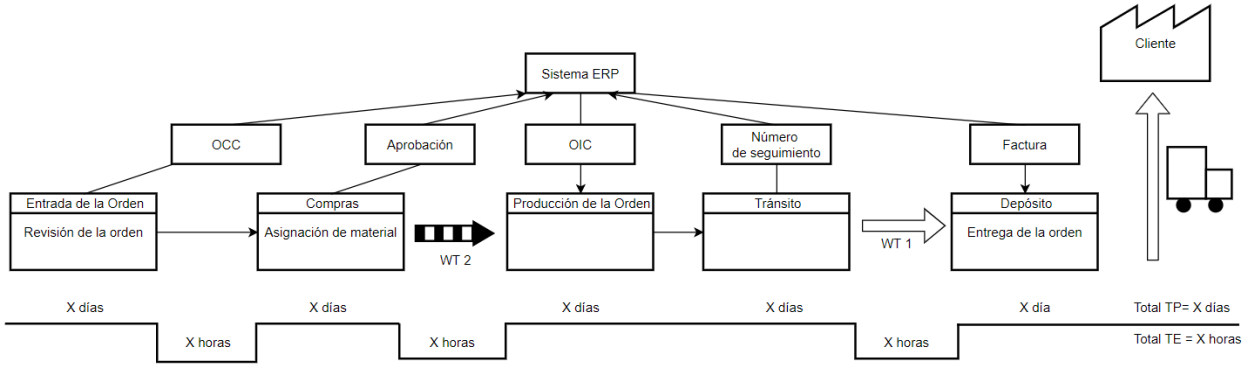


Figura 53 Nuevo mapa flujo de valor "ciclo de la orden". Fuente: Elaboración propia, 2021.

Nota: Los valores con una "x" fueron suprimidos a solicitud del patrocinador.

El mapa flujo de valor deja claro que el tiempo de espera representa un 39% del tiempo total, se acepta el periodo de compras como normal para el negocio y no como un agregado al tiempo de espera. Es muy importante aclarar el uso de los números promedio de la figura 53 versus a la diversidad de variables contenidas en una orden, lo cual significa que el tiempo total no siempre será de "x" días.

Capítulo 6 Conclusiones y recomendaciones

6.1 Conclusiones

Este proyecto fue dotado con el uso de grandes metodologías de trabajo comprobadas mundialmente; en primera instancia: la metodología *Lean Six Sigma*. En breve se detalla la distribución del cumplimiento de cada objetivo por cada etapa del DMAIC.

Tabla 32 Cumplimiento de objetivos utilizando DMAIC

<table>
<tr><th>Definir</th><th colspan="2">Medir</th><th>Análisis</th><th>Mejora</th></tr>
<tr><td rowspan="2">Capítulo #1
Sesiones:
1.3, 1.6, 1.7

Objetivo #1

Objetivo #2</td><td colspan="2">Objetivo #3
Objetivo #4</td><td rowspan="2">Objetivo #5</td></tr>
<tr><td colspan="2">Metodología CRISP-DM</td></tr>
<tr><td colspan="4">Objetivo general</td></tr>
</table>

Fuente: Elaboración propia, 2021.

Como se puede apreciar en la tabla 32, las etapas de medir y análisis fueron soportadas por la grandiosidad de la metódica CRISP-DM.

La incorporación de estas metodologías trajo muchos beneficios principalmente el enfoque de los objetivos del proyecto, por tratarse de “hojas de ruta” para encontrar soluciones y asegurar los mejores resultados. Al igual que agilizó la comprensión del tema y enfatizó el ítem/proceso de dolor para transformarlos en oportunidades de progreso.

En breve, se describen las deducciones de cada objetivo planteado en división 1.6.

6.1.1 Conclusión objetivo #1.

“Identificar las fuentes de datos que permitan la evaluación del riesgo o desvió de entregar a tiempo un pedido”. Este objetivo fue alcanzado gracias al uso de las herramientas del capítulo 4 (específicamente diagrama SIPOC) y se concluye: los sistemas ERP y almacén de datos actuales de la compañía no cuentan con el nivel de detalle histórico necesario para darle la correcta trazabilidad al ciclo de la orden en las etapas de interés para este proyecto.

A pesar de que se obtuvo la mayoría de las razones (variables) de una entrega tardía o no, las grandes limitantes fueron: calidad, precisión y completitud de los datos.

6.1.2 Conclusión objetivo #2.

“Distinguir los diferentes factores que intervienen en la entrega tardía de productos por parte del departamento de la cadena de suministros”. Este objetivo fue alcanzado y se concluye que: La visibilidad oportuna de declives operativos permitirá a los responsables actuar anticipadamente con planes de priorización o contingencia.

Los tiempos de espera están matando la efectividad del equipo SC: *Waiting Time 1 y 2*, por concepto. En el caso de la primera variable, las posibilidades de rescatar los tiempos de entrega son casi nulas por tratarse de un paso antes de facturación, no obstante, el estudio de las causas podría salvar el día y convertir está orden en dinero en menor tiempo del pensado, con los datos que tenemos y alcance de este proyecto, todo apunta a la presencia de órdenes condicionadas a entrega por número de cotización.

Varias de las hipótesis por parte del equipo SC cuentan con el respaldo suficiente de datos para aprobar o refutar la suposición, excelente indicio que confirma el alto nivel de experiencia en los participantes seleccionados, lastimosamente, no todas las hipótesis tuvieron la misma suerte, por causa de datos faltantes o discrepantes. Al respecto, se deducen tres teorías: el rango de tiempo seleccionado no fue el suficiente, el periodo de tiempo está afectado por estacionalidad llámese Navidad o Covid-19 y la tercera los factores no se presentan con tanta frecuencia o están siendo opacados/traslapados por los altos índices en días de espera y traslado de materia prima o producto final.

6.1.3 Conclusión objetivo #3.

“Desarrollar un modelo predictivo utilizando los factores encontrados en el objetivo 2”. Este objetivo fue alcanzado y se concluye: las métricas de los modelos alcanzaron valores positivos muy altos de lo esperado, demostrando la viabilidad de implementación en gran escala para la empresa patrocinadora.

La mayoría de las variables cualitativas no representaron la utilidad confiada por los expertos de negocio ni siquiera usándolas combinadas o solas, ejemplo: *Hold*.

La ausencia de la variable *Waiting Time* 1 en los métodos causaba un descenso en los resultados de las métricas, su necesidad indica que el negocio debe trabajar en esta área urgentemente.

6.1.4 Conclusión objetivo #4.

“Desglosar el ciclo de producción a días calendario”. Este objetivo fue alcanzado, ver división 5.2 y se concluye que los tiempos de espera pre y post

producción son factores que no deben tomarse a la ligera, están jugando un papel muy importante en la alteración o incumplimiento de los tiempos de entrega.

La condición de entrega por cotización requiere especial atención por la alta gerencia, la ejecución de un plan de mitigación adecuado puede acelerar el proceso exponencialmente para todas las órdenes implicadas.

Por la complejidad del negocio, se decide no utilizar un promedio general para cada uno de los componentes del ciclo de vida de la orden, según la etapa; es fundamental considerar las variables: clase del producto, localización y condición de entrega ganando una mayor precisión en el cálculo.

6.1.5 Conclusión objetivo #5.

"Proponer una nueva metodología en estimación de tiempos de entrega, promoviendo un cambio cultural y mejorando la relación entre los clientes internos y externos". Este objetivo fue alcanzado y se concluye que: A lo largo de este proyecto se demuestra la necesidad de cambio y ajustes para lograr reducir la incertidumbre favoreciendo las relaciones entre el departamento de la cadena de suministro y sus clientes internos y externos, los sistemas presentan menos discrepancias; es decir, menos trabajo justificativo para el equipo SC, inclusive se presentaron oportunidades para la correcta ejecución de nuestros valores, todo esto se desarrolló en capítulo 5.

6.1.6 Conclusión objetivo general.

"Estimar el cumplimiento de la entrega de producto al cliente final para una empresa privada utilizando aprendizaje automático". Este objetivo fue alcanzado y se concluye que la propuesta final utiliza datos retrospectivos para satisfacer dos ejes, el primero fue el desarrollo de un modelo predictivo y la segunda modificación en los

cálculos de las etapas de producción, los cuales, sin lugar a duda, facilitaron la tarea de los gerentes en sus planes de ejecución futuros con una base de medición concreta.

Los objetivos específicos cooperaron enormemente en la creación de esta propuesta integral beneficiando los temas tocados en la sección 1.4.

6.2 Recomendaciones

En esta sesión se enumeran recomendaciones desde el punto de vista técnico y operacional:

6.2.1 Recomendaciones técnicas.

Almacén de datos: La publicación de los cálculos utilizados en cada campo del almacén de datos facilitarán la tarea de selección al usuario a la hora de realizar sus análisis o proyectos, llamado técnicamente “diccionario de datos”.

Sistemas ERP: una excelente mejoría sería el llenado completo de las fechas reales de cada ciclo de la orden o bien perfeccionamiento de las señales entre las entidades que intervienen en el proceso y el sistema ERP.

Recategorización de los estatus de producción: Se aconseja segmentar las señales de producción actual a los estatus que maneja la fábrica, evitando la generalización y libre interpretación por parte de SC.

Características predictivas: Una acción con excelente impacto en los administradores del país sería la incorporación de en las herramientas/reportes con funciones de predicción; en este momento, los usuarios no lo piden porque desconocen de este gran poder tecnológico.

6.2.2 Recomendaciones operativas.

Alianza con compradores de materiales: La visibilidad en este tema ayudaría a SC a conocer con anticipación si su orden cuenta o no con el material requerido para construcción, o bien, cuando una orden toma los materiales de otra por conceptos de estrategia y negociación.

Reporte del estatus de las órdenes: es importante concienciar al equipo creador del reporte en la asignación correcta del estatus de la orden, orden lógico, cierre de los ciclos y la actualización a tiempo y explicarles los resultados futuros que su trabajo puede lograr.

Equipo de "x": se aconseja agregar a la lista de tareas del equipo "x" el estudio de los tiempos de espera, apoyándolos con métricas automatizadas y dotándolos de planes de mitigación o contingencia a estos hallazgos.

Monitoreo de efectividad: Replicar este proyecto es una buena práctica de estudio retrospectivo y prospectivo al mismo tiempo, pues se puede dar seguimiento a los esfuerzos operativos mientras que se van afinando los modelos para futuras predicciones. Inclusive se recomienda la exploración sobre otro tipo de variables.

Reevaluar conceptos: Una vez que se controlen los hallazgos de este proyecto, se sugiere reevaluar el concepto de "entrega tardía" y "desvío" por parte del negocio.

Por otro lado, sería importante crear nuevas reglas de negocio en el tratamiento y seguimiento de un pedido cuyo cumplimiento sea por número de cotización, todas las

órdenes deberían ser tratadas como una sola evitándose el incremento en los tiempos de espera.

Capítulo 7 Reflexiones Finales

Este proyecto contribuyó notablemente en minimizar la incertidumbre de la eficacia operativa, promoviendo estimaciones aterrizadas a la realidad con poca variabilidad a lo largo del ciclo de vida de la orden, para favorecer la relación compañía-cliente. Otro punto fue el conocimiento proactivo de las posibles interrupciones o declives de una orden, facultando al equipo SC reducir el riesgo a tiempo; asimismo, para la liberación de tiempo laboral en colaboradores relacionados con este proceso, además de aminorar la dependencia a otros equipos como tecnología de información, finanzas, operaciones, entre otros; sin olvidar, la colaboración a la cultura analítica de datos, la empresa patrocinadora cuenta con bancos de datos muy valiosos que todavía no han sido explotados como se debería.

El aprendizaje ha sido en ambos sentidos: negocio e investigadora, este proyecto demostró que la adaptación y uso de la minería de datos es ideal para identificar y comprender factores que impactan en los resultados de la pregunta inicial, que, por añadidura, repercuten en grandes ganancias para la corporación una vez analizadas y traducidas en estrategias o factores de éxito.

Como aporte final se concluye que la creación de una base de datos para estudios prospectivos no es tan sencillo como parece, pues la idea es obtener datos significantes para al negocio, sin embargo, es inevitable lidiar con temas como datos defectuosos o basura, un punto clave fue la colaboración activa de los expertos.

La adopción de un predictor de declives para empresas manufactureras no es imposible, su nivel de complejidad amerita mucho trabajo de investigación, desarrollo del manejo de la variabilidad y sobre todo personal capacitado en estadística, *machine*

learning, bases de datos, más, el apoyo incondicional de especialistas en cada área operativa; ante una idea tan ambiciosa, la inversión para la corporación es mucho menor en comparación con los beneficios que se obtendrán a largo plazo.

Capítulo 8 Trabajos a futuro

La propuesta de posibles trabajos o mejoras a futuro serían:

- Creación de un *datawarehouse* con los datos históricos a medida de lo requerido por el negocio/modelo dado que los existentes cuentan con lógica multipropósito para uso de diferentes equipos.
- Incrementar la capacidad del modelo para que calcule el tiempo aproximado de entrega en días, según la propuesta de variables sugerida en este proyecto desde el día uno, para tratar de imitar la capacidad de predicción del tiempo de entrega de compañías como “Uber. Adicionalmente, habilitar la escogencia del cálculo de las variables en días calendario o días hábiles.
- Aplicar este proyecto, utilizando aquellos periodos en donde los temas de estacionalidad no sean de alto impacto.
- Implementar la visualización del modelo en herramientas populares como: PowerBI, Tableu, Qliksense etc.
- Sistema de advertencias: Sería de una gran ayuda para el administrador de país, recibir alertas sobre las órdenes que proyecten declives en los tiempos esperados.
- Incluir en el modelo variables temas excluidos en la sección limitaciones como, por ejemplo: ciclo de vida del producto y compra de materia prima (¿Es compra anticipada?, estudio de compras por producto).
- Experimentar con otros métodos y modelos para mejorar el desempeño de las predicciones y optimización del proceso de aprendizaje.

- La segmentación oportuna de los datos facilitará el manejo de la variabilidad en el análisis e implementación, sugerencia por producto, condición de entrega y localización o bien limitar la escogencia de variables, todo esto con el cuidado de verificar la participación equitativa, según la segmentación seleccionada, tratando de evitar sesgo.

Referencias bibliográficas

Polim, R. (2016). Real-Time Supply Chain Analytics-Shipment Duration Prediction.

Kusrini, E., Sugito, E., Rahman, Z. M., Setiawan, T. N., & Hasibuan, R. P. (2021). Risk mitigation on product distribution and delay delivery: A case study in an Indonesian manufacturing company. In IOP Conference Series: Materials Science and Engineering (Vol. 722, No. 1, p. 012015). IOP Publishing.

Wang, T., Xing, Z., Hu, H., & Qu, X. (2019). Overbooking and delivery-delay-allowed strategies for container slot allocation. Transportation Research Part E: Logistics and Transportation Review, 122, 433-447.

Güller, M., Koc, E., Hegmanns, T., Henke, M., & Noche, B. (2015). A simulation-based decision support framework for real-time supply chain risk management. International Journal of Advanced Logistics, 4(1), 17-26.

Valero Berjaga, A. (2013). Development of a framework for the integration of real-time data in supply chain risk management using simulation tools (Bachelor's thesis, Universitat Politècnica de Catalunya).

Elangovan, D., Sundararaj, G., Devadasan, S. R., & Karuppuswamy, P. (2010). Analysis of time delays for loss reduction in an automobile components manufacturing organisation using a proactive supply chain risk management model. International Journal of Services, Economics and Management, 2(3-4), 385-411.

Mohamed, A. A. M., & Coutry, N. (2015). Analysis of lead time delays in supply chain: a case study. International Journal of Economics and Management Engineering, 9(6), 2065-2070.

Li, C., & Liu, S. (2013). A robust optimization approach to reduce the bullwhip effect of supply chains with vendor order placement lead time delays in an uncertain environment. Applied Mathematical Modelling, 37(3), 707-718.

Gudum, C. K. (2002). On the distribution of lead time delays in supply chains. Copenhagen Business School.

2021 Viguera Editores S.L.U. (2018, 16 junio). Uso del test de evaluación cognitiva de Montreal (MoCA) en América Latina: revisión sistemática: Neurología.com. Recuperado de https://www.neurologia.com/articulo/2017508

Alicia Martín-Navarro, María Paula Lechuga Sancho, José Aurelio Medina-Garrido. (2018, 25 enero). Vista de BPMS para la gestión: una revisión sistemática de la literatura | Revista Española de Documentación Científica. Recuperado de https://redc.revistas.csic.es/index.php/redc/article/view/1019/1620

Costa, M. A. B. (2016, 14 junio). Análise dos modelos e atividades do pré-desenvolvimento: revisão bibliográfica sistemática. Recuperado de https://www.scielo.br/j/gp/a/srcqKXTkLthVVwxNmFhddFB/?lang=en

Hessing, T. (2021, 22 junio). Cause and Effect Diagram (aka Ishikawa, Fishbone). Recuperado de https://sixsigmastudyguide.com/cause-effect-diagram-aka-ishikawa-fishbone-herringbone-fishikawa/

Microsoft. (2021, 1 enero). Calculadora de precios. Recuperado de https://azure.microsoft.com/es-es/pricing/calculator/

Mancilla-Vela, G. (2020, 1 diciembre). Factors associated to student success in online learning: a data mining analysis. Recuperado de

https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-50062020000600023&lng=en&nrm=iso&tlng=en

P. (2021, 15 septiembre). Value Stream Mapping o Mapa de Flujo de Valor». Recuperado de http://www.consultoriaprocesos.com/value-stream-mapping/

Meng Huang y Masood Bagheri (2019) Predicting Deviation in Supplier Lead Time and Truck Arrival Time Using Machine Learning, Master's thesis in Supply Chain Management, Chalmers University of Technology, Recuperado de https://odr.chalmers.se/bitstream/20.500.12380/257041/1/257041.pdf

Silvera, Yolanda (2017) Supply chain visibility tools usage and their impact on on-time delivery: a case study of a fast-moving consumer goods small, medium enterprise in London, Degree of Doctor of Philosophy, Brunel University London Recuperado de https://bura.brunel.ac.uk/handle/2438/15663

Silva, N., Ferreira, L. M. D., Silva, C., Magalhães, V., & Neto, P. (2017). Improving supply chain visibility with artificial neural networks. Procedia Manufacturing, 11, 2083-2090. Recuperado de https://www.sciencedirect.com/science/article/pii/S2351978917305371

Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., & McFarlane, D. (2021). Supply chain data analytics for predicting supplier disruptions: a case study in complex asset manufacturing. International Journal of Production Research, 58(11), 3330-3341. Recuperado de https://www.tandfonline.com/doi/abs/10.1080/00207543.2019.1685705

Barbara Kitchenham and Pearl Brereton, (2013) A Systematic Review of Systematic Review Process Research in Software Engineering, Keele University,

Recuperado de https://eprints.keele.ac.uk/2683/1/B%20Kitchenham%20-%20A%20systematic%20review%20of%20systematic%20review%20process%20research%20in%20software%20engineering.pdf

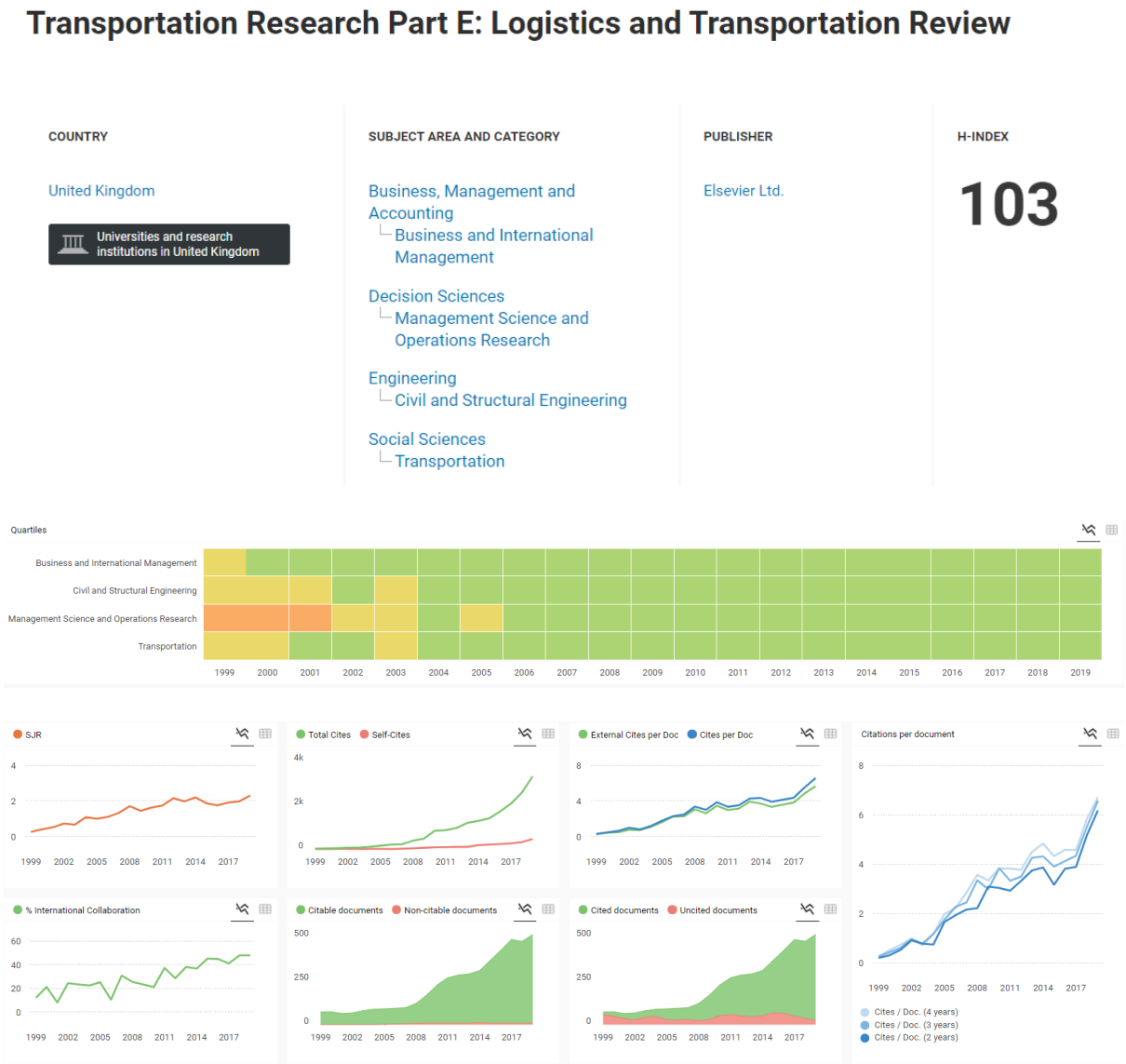
IBM. (1994). CRISP-DM Help Overview. Recuperado de https://www.ibm.com/docs/en/spss-modeler/SaaS?topic=dm-crisp-help-overview

Scimago Lab. (2010). Scimago Journal & Country Rank. Recuperado de https://www.scimagojr.com/

Apéndices

Evaluación de la calidad de los estudios

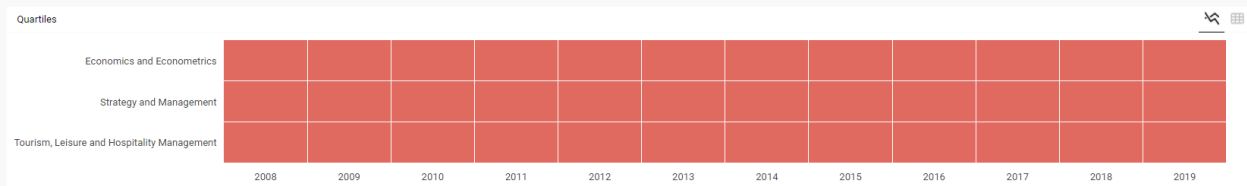
A continuación, se ilustran las calificaciones de cada estudio mencionados en la sección 1.9.5.3



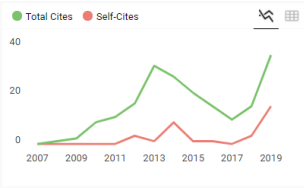
International Journal of Services, Economics and Management

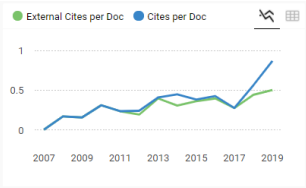
COUNTRY	SUBJECT AREA AND CATEGORY	PUBLISHER	H-INDEX
Switzerland Universities and research institutions in Switzerland	Business, Management and Accounting Strategy and Management Tourism, Leisure and Hospitality Management Economics, Econometrics and Finance Economics and Econometrics	Inderscience Publishers	9

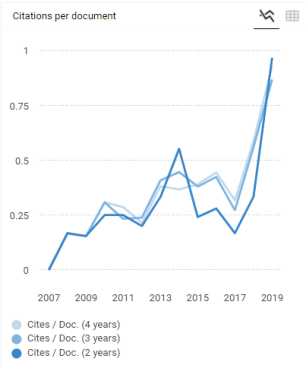
PUBLICATION TYPE	ISSN	COVERAGE	INFORMATION
Journals	17530822, 17530830	2014-2019	Homepage How to publish in this journal Contact



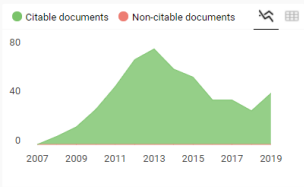


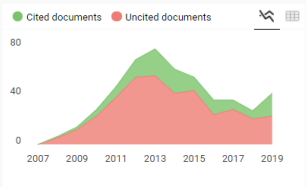








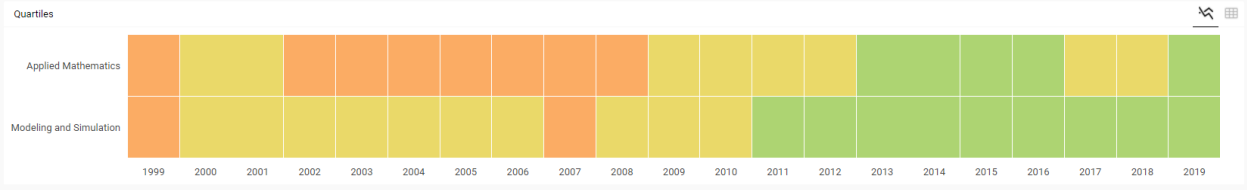


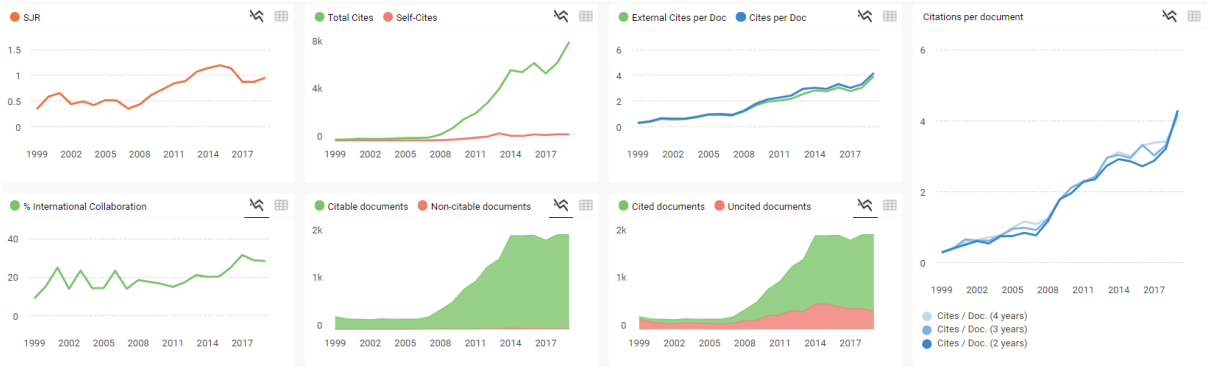


Applied Mathematical Modelling

COUNTRY	SUBJECT AREA AND CATEGORY	PUBLISHER	H-INDEX
Netherlands Universities and research institutions in Netherlands	Mathematics Applied Mathematics Modeling and Simulation	Elsevier Inc.	103

PUBLICATION TYPE	ISSN	COVERAGE	INFORMATION
Journals	0307904X	1976-2020	Homepage How to publish in this journal Contact





Comprobación por universidad

Rank	University	Overall Score	International Students Ratio	International Faculty Ratio
101	Pennsylvania State University University Park,United States	58.4	36.9	15.9

Rank	University	Citations per Faculty	Academic Reputation	Employer Reputat
101	Pennsylvania State University University Park,United States	59.7	74.9	75.3

Rank	University	Overall Score	International Students Ratio	International Faculty Ratio
314	Universitat Politècnica de Catalunya ·... Barcelona,Spain	33.2	26.3	11

Rank	University	Citations per Faculty	Academic Reputation	Employer Reputat
314	Universitat Politècnica de Catalunya ·... Barcelona,Spain	21.6	33.7	32.3

Rank	University	Overall Score	International Students Ratio	International Faculty Ratio
N/A	Copenhagen Business School Frederiksberg,Denmark		65.5	85.9

Rank	University	Citations per Faculty	Academic Reputation	Employer Reputat
N/A	Copenhagen Business School Frederiksberg,Denmark	41.7		19.1