



Desarrollo de una alternativa de automatización para el procesamiento de direcciones en el sistema última milla del Grupo Familia

Maria Paulina Vélez Vélez

Informe de práctica para optar al título de Ingeniera Industrial

Asesor

Miguel Ángel Arroyave Guerrero, M. Sc. Ingeniería.

Universidad de Antioquia
Facultad de Ingeniería
Ingeniería Industrial
Medellín, Antioquia, Colombia
2022

Cita	(Vélez Vélez, 2022)
Referencia	Vélez Vélez, M. (2022). <i>Desarrollo de una alternativa de automatización para el procesamiento de direcciones en el sistema última milla del Grupo Familia</i> . [Trabajo de grado profesional]. Universidad de Antioquia, Medellín, Colombia.
Estilo APA 7 (2020)	



Grupo Familia – Desarrollo de Negocios Logísticos



Centro de Documentación Ingeniería (CENDOI)

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano/Director: Jesús Francisco Vargas Bonilla.

Jefe departamento: Mario Alberto Gaviria Giraldo.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Tabla de contenido

Resumen	7
Abstract	8
Introducción	9
1. Objetivos	11
1.1 Objetivo general	11
1.2 Objetivos específicos	11
2. Marco teórico	12
3. Metodología	15
3.1 Reconocimiento del proceso	15
3.1.1 Entrevistas	15
3.1.2 Exploración del proceso	15
3.1.3 Puntos críticos	15
3.2 Priorización de alternativas	16
3.2.1 Selección de alternativas	16
3.2.2 Búsqueda del estado del arte	16
3.3 Selección de la propuesta	16
3.3.1 Propuesta para la estructuración de datos	16
3.3.2 Propuesta para la transformación de datos	17
3.4 Comparación de métodos y toma de tiempos	17
3.4.1 Disminución de tiempos de ejecución	17
3.4.2 Respuesta del equipo de trabajo:	17
3.5 Tabla resumen	17
4. Resultados	19
4.1 Reconocimiento del proceso:	19

4.1.1 Entrevistas	19
4.1.2 Exploración del proceso	21
4.1.3 Puntos críticos	22
4.2 Priorización de alternativas	24
4.2.1 Selección de alternativa	24
4.2.2 Búsqueda del estado del arte	25
4.3 Selección de la propuesta	26
4.3.1 Propuesta para la estructuración de datos	26
4.3.2 Propuesta para la transformación de datos	28
4.4 <i>Comparación de métodos y toma de tiempos</i>	31
4.3.1 Comparación de métodos y toma de tiempos	31
4.3.2 Respuesta del equipo de trabajo:	33
5. Conclusiones	36
Referencias	38
Anexos	40

Lista de Tablas

Tabla 1. Relación entre los objetivos específicos y las etapas de la metodología.....	17
Tabla 2. Descripción del proceso	19
Tabla 3. Ejemplo Limpieza de datos	20
Tabla 4. Toma de tiempo y anotaciones.....	21
Tabla 5. Alternativas de solución para la problemática	24
Tabla 6. Ventajas y desventajas de las alternativas de solución propuestas	25
Tabla 7. Diccionario de términos y abreviaturas para la dirección	27
Tabla 8. Diccionario de términos y abreviaturas para el complemento	27
Tabla 9. Aplicación metodología Expresiones Regulares.....	28
Tabla 10. Estructura del código en VBA.	29
Tabla 11. Comparación de tiempos entre métodos	32

Lista de Figuras

Figura 1. Puntos críticos del proceso.	22
Figura 2. Mejoramiento del proceso actual.....	33
Figura 3. Funcionalidad y eficiencia del proceso.....	34
Figura 4. Entendimiento del proceso.....	35
Figura 5. Satisfacción de la propuesta.....	35

Resumen

Entre las consecuencias generadas por el COVID-19 se encuentran los cambios en la logística de entrega que obligaron a muchas empresas a migrar al sistema de última milla. Este proceso consiste en realizar una limpieza de datos, georreferenciar y crear una ruta óptima para generar las hojas de ruta, donde la mayoría de estas actividades se llevan a cabo de manera manual y por ende, se invierten grandes períodos de tiempo en su ejecución. Una alternativa para automatizar este proceso es la creación de una macro de Excel con lenguaje de Expresiones Regulares, la cual permite la transformación del campo de dirección digitada por los usuarios en dos únicos datos: Uno asociado a la dirección estructurada y otro al complemento, los cuales sirven para ser georreferenciados más adelante. Con la implementación de esta alternativa de automatización en una base de datos de 450 usuarios, se encontró una disminución del 82,59% del tiempo invertido en el proceso de limpieza de datos y georreferenciación. Para la limpieza de datos se obtuvo un tiempo operativo de 726,41 segundos comparado con el tiempo asociado del método inicial, que era de 13536 segundos con una eficiencia del 94,67%. A su vez, el proceso de georreferenciación pasó de llevarse a cabo en 13954,5 segundos a hacerlo en 567,08 segundos con una eficiencia del 98,3%.

Palabras clave: automatización, limpieza de datos, georreferenciación, macro, expresiones regulares, sistema última milla, procesamiento de direcciones, eficiencia.

Abstract

Among the consequences generated by COVID-19 are the changes in delivery logistics that forced many companies to migrate to the last mile system. This process consists in performing data cleaning, georeferencing and creating an optimal route to generate the roadmaps, where most of these activities are carried out manually and therefore, large periods of time are invested in their execution. An alternative to automate this process is the creation of an Excel macro with Regular Expressions language that allows the transformation of the address field typed by users into two unique data: One associated with the structured address and another with the complement, which be of service to be georeferenced later. With the implementation of this automation alternative in a database of 450 users, an 82.59% decrease in the time spent in the data cleaning and georeferencing process was found. For data cleaning, an operating time of 726.41 seconds was obtained compared to the associated time of the initial method, which was 13536 seconds with an efficiency of 94.67%. Simultaneously, the georeferencing process went from being carried out in 13954.5 seconds to doing it in 567.08 seconds with an efficiency of 98.3%.

Keywords: automation, data cleaning, georeferencing, macro, regular expressions, last mile system, address processing, efficiency.

Introducción

Uno de los cambios más significativos que dejó la pandemia del COVID 19, fue la necesidad de las empresas para hacer llegar los productos sin intermediación hacia el consumidor final, proceso denominado última milla.

El sistema de última milla consiste en la entrega de paquetes de bajo volumen al cliente, es un proceso que le agrega valor a la cadena de suministro, y, según Francisco Aranda, presidente de UNO, en el 70% de los casos, el consumidor de comercio electrónico reconoce que repite la compra en un portal si el proceso de entrega ha sido satisfactorio. (Gonzalez, 2019)

Por esta razón, no sólo se requiere llegar al consumidor, sino también hacerlo de forma oportuna para realizar una entrega satisfactoria; el input principal para realizar la entrega consiste en la calidad de la información suministrada por parte de los consumidores, tanto de los productos que requiere, como de la dirección para el envío.

El Grupo Familia, compañía líder en higiene y salud a nivel global (Somos Grupo Familia, n.d.) lleva alrededor de dos años implementando el sistema de última milla para sus entregas, sin embargo, ha tenido una complicación con la calidad de los datos obtenidos en las direcciones, ya que estos no llegan con la estructura deseada.

Se plantearon una serie de alternativas que darían solución a las inconsistencias encontradas en el registro de las direcciones, entre ellas se encuentran la implementación de la inteligencia artificial, la adecuación del formulario donde se realiza la inscripción por el usuario y la aplicación de las herramientas de la minería de datos.

Con la implementación de una macro de Excel con Expresiones Regulares para normalización de direcciones se logró reducir un 94,66% los tiempos de ejecución con una eficiencia del 94,67%. Además, con la extensión de Google Drive, Geocode by Awesome Table, los tiempos de georeferenciación disminuyeron un 95,94% con una eficiencia del 98,3%. Esta

propuesta fue bien recibida por el equipo de trabajo del Grupo Familia, ya que la solución planteada automatiza el proceso y disminuye tanto los tiempos de ejecución como los costos del mismo.

1. Objetivos

1.1 Objetivo general

Proponer una alternativa de automatización para el procesamiento de direcciones en el Grupo Familia para disminuir los tiempos de ejecución en georreferenciación del sistema última milla.

1.2 Objetivos específicos

- Realizar un diagnóstico del proceso actual para identificar la estructura, la serie de actividades realizadas y los puntos clave de intervención.
- Generar y priorizar alternativas de solución que permitan procesar las direcciones de modo eficiente.
- Proponer una alternativa para la automatización del procesamiento de direcciones del sistema última milla en el Grupo Familia enfocado en la limpieza de datos.
- Medir los tiempos de ejecución en el proceso de georreferenciación del sistema última milla, con el fin de determinar la eficiencia y veracidad de la propuesta planteada frente al método actual.

2. Marco teórico

A raíz de la pandemia y los cambios que esta ha supuesto en los procesos dentro y fuera de las empresas, el concepto de última milla ha venido tomando fuerza. La necesidad de los clientes de recibir el producto en la puerta de sus casas a causa de las restricciones de movilidad supuso todo un reto en temas de logística, pues el consumidor quiere tener cada vez más control sobre cuándo, cómo y dónde recibe los productos que ha comprado, por lo que se hace necesario disminuir los tiempos de espera para ajustarse a las demandas de los clientes. (Cortés, 2021)

Para el caso del Grupo Familia, compañía líder en higiene y salud a nivel global (Somos Grupo Familia, n.d.) se ha impulsado el sistema de última milla, con el propósito de hacer llegar sus productos al consumidor final sin intermediación. En el camino, se ha observado que la digitación de las direcciones por parte de los usuarios, genera datos confusos, los cuales deben someterse a un proceso de limpieza de datos para continuar la entrega del pedido. El proceso de análisis y procesamiento de estas direcciones resulta complejo y extenso dado al gran número de datos almacenados y la falta de estructura de estos. (Troncoso y Fernández, 2021)

Este proceso requiere la modificación manual, lo que significa una demora en el proceso de entrega de pedidos, lo que significa más tiempo de espera por parte del cliente. El proceso de limpieza de datos sirve para determinar datos inexactos, incompletos o ilógicos y luego mejorar su calidad mediante la corrección de errores y omisiones que se detecten; este proceso puede incluir un formato de control, controles de integridad, razonabilidad y límites y la capacidad de revisar los datos para identificar valores atípicos. (Chapman, 2005)

Estas direcciones geocodificadas son el insumo para la georreferenciación, la cual permite determinar la posición de un punto de entrega en un sistema de coordenadas espacial diferente al que se encuentra. Este tipo de referenciación se utiliza con frecuencia en los sistemas de información geográfica (SIG). (Dávila Martínez & Camacho Arranz, 2012)

Realizar una geo codificación manual a bases de datos de gran volumen requiere mucho tiempo y resulta poco eficiente. Un estudio realizado a la base de datos de direcciones donde se

han presentado hechos delictivos, proporcionada por la unidad de análisis penal de la Fiscalía Regional de Biobío, Chile, evidencia que los tiempos de ejecución disminuyen un 94% al automatizar el proceso de geocodificación. Adicionalmente, se reduce un 25.22% la necesidad de georreferenciar manualmente las localizaciones. (Troncoso y Fernández, 2021)

El principal problema que afecta el procesamiento de direcciones de Última milla es la normalización de las direcciones en las bases de datos suministradas por las diferentes marcas. La normalización de direcciones es una actividad fundamental para la logística de entregas en una empresa, ya que esta permite la aplicación de normas para verificar las direcciones según los ficheros suministrados, para generar datos limpios, estandarizados y validados. Además, esta normalización trae beneficios como la reducción de costos, los tiempos de ejecución y de entrega. (*Importancia de normalización de direcciones en BBDD*, 2021)

Existen diferentes alternativas para dar solución a la falta de normalización en las direcciones, como la implementación de inteligencia artificial, adecuar el formulario que debe llenar el usuario, la aplicación de herramientas de minería de datos y la adecuación de una macro de Excel VBA.

Al indagar con las diferentes marcas de la compañía, se llegó a la conclusión de que entre más campos diferentes deba digitar el usuario, menor será su interés en finalizar la inscripción, por lo que en muchos casos se termina perdiendo la atención del cliente. Teniendo en cuenta lo anterior y considerando que el alcance del proyecto no permite la modificación de los formularios, se descarta esta alternativa.

Por otro lado, la inteligencia artificial ofrece soluciones a problemas planteados, mediante integración algorítmica y automatización de procesos, simulando la inteligencia humana por un computador (Martinez Izquierdo, 2009), sin embargo, esta alternativa implica un amplio desarrollo e inversión de tiempo por lo tanto, no es factible su aplicación dado el alcance del trabajo.

El proceso de normalizar los parámetros de las direcciones se realiza de manera manual, por lo tanto, se podría optar por la implementación de la minería de datos. Esta es una herramienta

ampliamente difundida en la actualidad, ya que se encarga de preparar, sondear y explorar los datos para sacar la información oculta y útil en ellos, (Monjas Belinchón, sf.). Una herramienta de la minería de datos son las expresiones regulares, mediante las cuales es posible hacer búsquedas contextuales en las direcciones ingresadas por los usuarios y realizar modificaciones sobre las mismas, incrementando la eficiencia del proceso y con esto sanear la problemática de normalización de direcciones.

Por otro lado, se plantea la implementación de Geocode de Google, herramienta que permite convertir direcciones en coordenadas geográficas, que son utilizadas para ser marcadas en el mapa y programar los envíos a los usuarios (Google Developers). Esta herramienta posibilita la georreferenciación masiva de direcciones en tiempo reducido.

Para aplicar la herramienta expresiones regulares, se desarrolla una macro en Visual Basic for Application de Excel, la cual incluye dos diccionarios para estandarización de términos y a la segmentación de los mismos, logrando la normalización de la información.

3. Metodología

El presente trabajo se desarrolló mediante cuatro etapas metodológicas con las cuales se logró el reconocimiento del procesamiento de direcciones del sistema última milla, se llevó a cabo una búsqueda de soluciones lo que permitió la selección de una alternativa eficiente de automatización para la disminución de tiempos de ejecución, y la medición de los resultados obtenidos con el desarrollo de la propuesta. Estas etapas metodológicas se especifican a continuación:

3.1 Reconocimiento del proceso

Esta etapa consistió en comprender el paso a paso llevado a cabo para los procesos de última milla, incluyendo la limpieza de datos, estandarización de la estructura de direcciones, georreferenciación y ruteo. Para lograr este resultado se realizaron entrevistas, análisis de las fuentes de la información, bases de datos y visualización de procesos terminados que evidenciaban las falencias del proceso actual.

3.1.1 Entrevistas: inicialmente, se realizó una serie de entrevistas a los integrantes del equipo, con el propósito de reconocer la metodología empleada para el procesamiento de bases de datos suministradas por los diferentes clientes internos de la compañía, con la finalidad de impactar al cliente final mediante paqueteo por el sistema de última milla. Una vez comprendido el proceso, se elaboró un diagrama de flujo que permitió ver la secuencia de actividades realizadas en dicho proceso.

3.1.2 Exploración del proceso: se realizó una prueba con una base de datos remitida por la marca Nosotras, la cual se procesó con la metodología actual: limpieza y estructuración de datos, separación de dirección y complemento, georreferenciación de cada punto y construcción de un ruteo (VRP), para la posterior programación de transporte.

3.1.3 Puntos críticos: se realizó un diagrama de Ishikawa para identificar los principales problemas y las respectivas causas asociadas a las diferentes categorías.

3.2 Priorización de alternativas

La segunda etapa consistió en el planteamiento de soluciones que permitieran la automatización del proceso actual, teniendo como objetivo la disminución de tiempos de ejecución en el proceso de última milla. Basados en los puntos críticos se descartaron las alternativas enfocadas en los factores externos sobre los cuales no se tiene control y desde allí se realizó una búsqueda exhaustiva del estado del arte.

3.2.1 Selección de alternativas: una vez identificados los puntos críticos, se plantearon alternativas ante cada una de las causas que originaban la problemática actual, sin embargo, surgían soluciones que se salían del alcance del proyecto y otras con poca viabilidad, por lo tanto no fueron tenidas en cuenta. Aquellas alternativas que continuaban en lista, apuntaban principalmente a la transformación masiva de la información.

3.2.2 Búsqueda del estado del arte: se hizo una exploración de alternativas existentes para problemáticas similares, las cuales apuntan a soluciones basadas en minería de datos, expresiones regulares, macros, inteligencia artificial, entre otras. Cada alternativa, se visualizó y se realizó una comparación, determinando ventajas y desventajas de su aplicación.

3.3 Selección de la propuesta

La tercera etapa permitió la elección de la propuesta, realizando el planteamiento de un proceso estandarizado, eficiente y ágil que disminuye los tiempos de ejecución considerablemente, además permite la estructuración estandarizada de la información contenida en las direcciones.

3.3.1 Propuesta para la estructuración de datos: se plantearon dos diccionarios que permiten la estandarización de términos comúnmente usados en las direcciones con sus diversas formas de escritura, lo cual permite que si un usuario ingresa una de las n formas de digitar “calle” este automáticamente la transforme al formato estándar “CL”.

3.3.2 Propuesta para la transformación de datos: se construyó una macro en Microsoft Excel VBA, con la metodología de expresiones regulares que permitió reconocer los caracteres pertenecientes a la dirección estructurada y el adicional separarlo en otra columna como el complemento de la misma, esto con ayuda de los diccionarios anteriormente creados.

3.4 Comparación de métodos y toma de tiempos

La última etapa consistió en comparar los métodos en aspectos de tiempo y calidad de la información. Además, se presentó la macro a los integrantes del equipo para validar su funcionalidad y aceptación.

3.4.1 Disminución de tiempos de ejecución: se llevó a cabo una prueba piloto con 450 direcciones para determinar el tiempo de ejecución y el impacto del nuevo método sobre los puntos críticos identificados.

3.4.2 Respuesta del equipo de trabajo: se realizó una encuesta a los integrantes del equipo de trabajo, con el objetivo de conocer la percepción acerca de la propuesta planteada.

3.5 Tabla resumen

En la **Tabla 1**, se presenta la relación entre los objetivos específicos y las etapas de la metodología, exponiendo los resultados obtenidos a partir del desarrollo de la propuesta de mejora.

Tabla 1. Relación entre los objetivos específicos y las etapas de la metodología

Objetivo específico	Etapas metodológicas	Resultados esperados
1. Diagnosticar el proceso actual y sus puntos clave de intervención	E1.1: Entrevistas	Comprensión del paso a paso llevado a cabo en el procesamiento actual de direcciones y los principales puntos a intervenir desde las diferentes perspectivas de los entrevistados. Además, elaboración del diagrama de flujo que presenta la secuencia de actividades.
	E2.2: Exploración del proceso	Reconocimiento propio del proceso actual llevado a cabo

	E3.3: Puntos críticos	Identificación de los principales puntos críticos y su impacto en el proceso.
2. Generar y priorizar alternativas de solución	E2.1: Selección de alternativas	Generación de diversas alternativas de solución ante los puntos críticos con posibilidad de intervención
	E2.2: Búsqueda del estado del arte	Reconocimiento de soluciones realizadas por otras personas en problemáticas similares
3. Desarrollar propuesta de mejora	E3.1: Propuesta para la estructuración de datos	Planteamiento y desarrolló de un diccionario para la transformación de términos comúnmente conocidos en las direcciones por un formato estándar.
	E3.2: Propuesta para la transformación de datos	Implementación de un código basado en expresiones regulares y el diccionario que permita el procesamiento masivo de direcciones en tiempos cortos.
4. Evaluar impacto y tiempos de la propuesta	E4.1: Disminución de tiempos de ejecución	Cálculo e interpretación de indicadores de mejora que comparen el método actual y el desarrollado valorando el tiempo y la eficiencia.
	E4.2: Respuesta del equipo de trabajo	Conocimiento de la percepción de los integrantes del equipo sobre el macro planteada.

Fuente. Elaboración propia.

4. Resultados

Con base en el análisis realizado sobre el procesamiento de direcciones se obtuvieron una serie de resultados que aportaron a la mejora del proceso gracias a la automatización por medio de herramientas tecnológicas, disminuyendo los tiempos de ejecución y aumentando la eficiencia. A continuación, se describen los resultados del proyecto.

4.1 Reconocimiento del proceso:

El proceso de entregas bajo la modalidad última milla, es la transformación que optó el Grupo Familia, para seguir llegando a sus clientes en tiempos de pandemia, y el cual ha perdurado en el tiempo, gracias a su buen funcionamiento y satisfacción del cliente. Sin embargo, actualmente las plataformas de registro en la cuales se captan los datos principales del cliente, cuentan con falencias, dado que da la libertad al usuario de escribir su información de cualquier forma. La consecuencia de esto, es una posterior y estricta transformación de la información para lograr las entregas al cliente. Esta metodología, requiere de procesamiento manual y horas hombre que hacen que se torne poco eficiente y tedioso. El proceso llevado a cabo y los puntos críticos asociados se especifican a continuación.

4.1.1 Entrevistas:

Mediante la entrevista a cada miembro del equipo, se logra obtener información cualitativa importante para la descripción y conocimiento del proceso. En la **Tabla 1** **Tabla 2**, se presenta un resumen de las principales características y conceptos metodológicos identificados.

Tabla 2. Descripción del proceso

Actividad	Descripción	Fases
Limpieza de datos	Estructurar adecuadamente las direcciones proporcionadas por los clientes con el fin de georreferenciar de manera precisa cada localización.	Identificar municipio
		Estructurar nomenclatura
		Separar complemento
Georreferenciación		Identificación de longitudes y latitudes

	Determinar la ubicación geográfica exacta de los puntos o las intersecciones que se deben visitar.	Alimentar el software de optimización de rutas
Optimización:	Determinar las rutas óptimas que se deben seguir para la entrega efectiva de las muestras.	Establecer criterios y requerimientos del proyecto: función objetivo, capacidades, disponibilidades, costos, etc.
Estructuración de las hojas de ruta	Compartir con los conductores el orden de las entregas y los datos personales de cada cliente que deben visitar, según el resultado de la optimización.	Realizar hoja de ruta según los resultados de la optimización.
Seguimiento a las entregas	Identificar las entregas exitosas y las que presentaron novedades para reintegrarlas al modelo logístico y programarlas nuevamente.	Registro documental de las entregas. Comunicación interdisciplinaria para hacer efectiva la entrega.

Fuente. Elaboración propia.

Luego de unificar las respuestas a la entrevista, se construye el diagrama de flujo para el proceso, ver **Anexo 1**. Inicialmente se hace recepción de la información entregada por cada marca sobre la última milla; si la información está completa, se procede a realizar revisión manual de cada dirección; posteriormente se realiza una reestructuración de la misma siguiendo el método actual preestablecido: dirección estructurada y complemento; una vez todas las direcciones están correctamente codificadas, se procede a georreferenciarlas en Google Maps y se crea una VRP para cada una y como paso final, se programan los envíos.

En la **Tabla 3**, se evidencia un ejemplo del resultado obtenido una vez se realiza el proceso de limpieza de datos en las direcciones para entregas. En este caso la dirección ingresada por el usuario fue Calle 134 Sur N 34 25, Apto 402 edf El Portal.

Tabla 3. Ejemplo Limpieza de datos

Dirección	Complemento
CL 134 sur #34-25	Edificio El Portal, Apto 402

Fuente. Elaboración propia.

4.1.2 Exploración del proceso:

Se realizó el proceso experimental con una muestra de 450 datos emitido por la marca Nosotras, en la cual se logró evidenciar las falencias que tenía el proceso. Para esto, se llevó a cabo cada una de las actividades explicadas en la **Tabla 2**, siguiendo la secuencia de pasos expuesta en el diagrama de flujo del proceso (ver **Anexo 1**), Obteniendo los resultados enunciados en la **Tabla 4**

Tabla 4. Toma de tiempo y anotaciones

Actividad	Tiempo de ejecución	Anotaciones importantes
Limpieza de datos	30,08 segundos modificar una dirección	Tiempos elevados para la modificación de datos, lo que implica horas-hombre y aumenta el porcentaje de error al ser un proceso manual.
	13536 segundos modificar 450 direcciones	
	225,6 minutos (3 horas 46 minutos) modificar 450 direcciones	
Georreferenciación	31,01 segundos georreferenciar una dirección	Tiempos elevados y proceso manual en la georreferenciación lo que aumenta las horas-hombre y el porcentaje de error.
	13954,5 segundos (232,58 minutos - 3 horas 53 minutos) georreferenciar 450 direcciones	
Optimización:	1230 segundos (20 minutos 50 segundos) para optimizar las rutas de 450 puntos	Tiempo aceptable, proceso estandarizado y dependiente del número de datos a iterar.
Estructuración de las hojas de ruta	10 minutos de socialización con los involucrados	Tiempo prudente, dado que depende del entendimiento de cada operador frente al proceso
Seguimiento a las entregas	20 minutos diarios mientras la operación esté vigente	Tiempo necesario, y formatos estandarizados para llevar control sobre las entregas procesadas.

Fuente. Elaboración propia

A partir de la experiencia, se pudo establecer, la absoluta necesidad de intervención para las actividades de limpieza de datos y georreferenciación, pues no sólo es un proceso que se realiza manualmente, sino que se hace una a una y entre más grande sea la base de datos, su tiempo de ejecución aumentará; además al realizarse manualmente aumenta el indicador de costos del operador que las transforma y el porcentaje de error.

Dadas estas implicaciones, se decidió validar dentro de ambas actividades, para determinar cuál de estas requiere mayor atención, en lo que se concluye que se debe atacar principalmente la “Limpieza de datos”, pues de allí parte el insumo principal para poder desempeñar las diferentes actividades del proceso actual.

Hacer “Limpieza de datos” consiste en separar la dirección del complemento, estructurar la dirección e identificar de qué municipio es la dirección; cada una de estas tareas es desempeñada de forma manual, tardando más del tiempo esperado en su transformación. A su vez, se indaga sobre el origen de los datos, en donde se reconoce que la información es captada mediante los formularios de inscripción que tiene cada una de las marcas, donde esperan los registros de los clientes para captar sus pedidos y posteriormente enviarlos hasta su destino.

4.1.3 Puntos críticos: con el propósito de determinar los puntos críticos que afectan en el sistema de última milla, se elaboró un diagrama de Ishikawa enfocado en la problemática “Limpieza de datos” haciendo énfasis en el procesamiento de direcciones que se puede observar en la **Figura 1**.

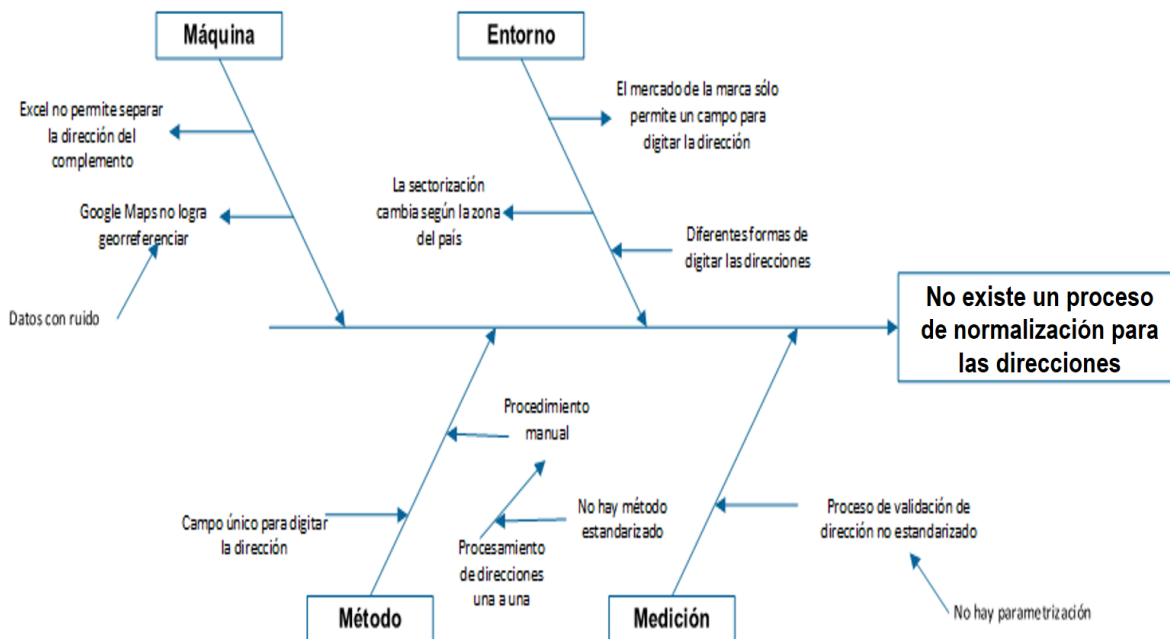


Figura 1. Puntos críticos del proceso.

Nota. Fuente elaboración propia.

En la actualidad, el proceso de entregas última milla presenta inconvenientes en la trazabilidad de la información en los formatos de inscripción, el problema principal radica en que el usuario ingresa los datos personales en la plataforma con el formato que desea generando diferencias significativas frente a lo deseado para darle continuidad al proceso.

Ante esta problemática se encuentran diferentes causas asociadas, desde la creación de los campos del formulario, hasta las herramientas con que posteriormente son transformados. Inicialmente se encuentran causas asociadas a la maquinaria, principalmente que no hay una fórmula de Excel que permita la diferenciación y separación de la dirección estructurada y el complemento. En cuanto a Google Maps que es la herramienta usada para georreferenciar, no es posible encontrar la ubicación, si se digita la dirección estructurada, junto al complemento.

También se identifican causas asociadas al entorno, entre ellas se encuentra que en las diferentes zonas del país se denominan las direcciones de forma diferente, su estructura cambia, así como sus componentes, mientras que en un departamento las direcciones se componen de “Calle”, “Carrera”, en otros constan de “Manzana”, “Lote”, lo cual dificulta la estandarización de formatos iniciales para las plataformas. Otra causa directamente relacionada con esta, es que los usuarios tienen infinitas formas de digitar la dirección de sus viviendas, dado que no hay una cultura que defina qué se debe digitar de determinada forma. Otra causal asociada al entorno, es que desde el mercadeo, las marcas han realizado estudios en los que concluyeron que entre más campos tenga que llenar un usuario en un formulario, menos deseos tendrá de continuar la inscripción.

En cuanto al método usado para transformar las direcciones, se establece que la causal principal es el impacto que tiene un único campo para digitar la dirección en los formularios de inscripción, pues así se le da la libertad al usuario para escribir su información sin seguir ningún parámetro. Además, el procesamiento posterior a la inscripción, es decir, el asociado a la transformación de los datos digitados por los usuarios, se realiza de forma manual, requiriendo así de mucho tiempo para procesarlas.

Finalmente, las causas asociadas a la medición, tienen que ver con la falta de estandarización y control tanto en la parte de inscripciones como en la posterior transformación de los datos.

De acuerdo con la información recolectada tanto en las entrevistas como en la exploración del proceso, se reconoce que se debe intervenir directamente en las causas asociadas al método, ya que de este modo se pueden automatizar los procedimientos manuales y así disminuir los tiempos de ejecución del proceso.

4.2 Priorización de alternativas

En la segunda etapa se realizó un análisis a cada una de las causas encontradas, determinando que las causas asociadas al entorno no tienen alcance de solución pues son factores incontrolables como lo es cambiar la nomenclatura de las direcciones de todo el país. A su vez, se realizó una búsqueda de información bibliográfica en la que se evidenciaron soluciones mediante: Minería de datos, Expresiones regulares, Macros, Inteligencia artificial, entre otras.

4.2.1 Selección de alternativa: a partir de las causas encontradas, se propusieron diversas soluciones enunciadas en la **Tabla 5** las cuales permiten darle fin a la problemática inicial. Sin embargo, se da una calificación a cada una de ellas donde se determina la viabilidad de su aplicación. Entre las cuales se descartan aquellas causas relacionadas con el entorno, puesto que están dentro de un escenario sobre el cual no se tiene control, por lo que no se puede modificar, especialmente por la complejidad de los estudios de mercadeo y la estructura existente de sectorización del país.

Tabla 5. Alternativas de solución para la problemática

Alternativa	Causa a impactar	Solución esperada	¿Es viable?
Reestructurar el formato del campo de Dirección desde los formularios de las marcas	Campo único para digitar las direcciones <hr/> Excel no permite separar las direcciones del complemento	Generar dos campos para que el usuario ingrese en uno la dirección estructurada y en el otro el complemento de la misma, así minimizar los tiempos de ejecución asociados a la separación de estas.	NO

Estandarizar los términos de las direcciones para crear un formato final único	No hay estandarización del método	Crear un diccionario que permita la iteración de términos semejantes y convertirlos en un formato estándar para estructurar tanto la dirección como el complemento.	SI
	La sectorización del país hace que cambien los términos en las direcciones		
Realizar una macro que permita la transformación y separación de las direcciones	Excel no permite separar la dirección del complemento	Desarrollar una macro que permita separar la dirección del complemento y de esta forma generar el input para la georreferenciación en Google Maps	SI
	Google Maps no georreferencia las direcciones de forma conjunta con el complemento		
Crear un modelo de inteligencia artificial	Excel no permite separar la dirección del complemento	Desarrollar una macro que permita separar la dirección del complemento y de esta forma generar el input para la georreferenciación en Google Maps	SI
	Google Maps no georreferencia las direcciones de forma conjunta con el complemento		

Fuente. Elaboración propia.

4.2.2 Búsqueda del estado del arte: ante la viabilidad de las propuestas anteriormente mencionadas, se realizó una búsqueda exhaustiva en los diferentes portales validando aplicaciones realizadas a problemas similares y reconociendo las ventajas y las desventajas de la aplicación de cada una de estas, ver **Tabla 6**. Este análisis, permitió descartar la alternativa de “Crear un modelo de inteligencia artificial” pues, aunque su aplicación permitiría un margen de error menor y unas mejoras significativas en el proceso, implicaría tanto inversión en tiempo como en capacitación para quienes seguirán usando el modelo, por lo que se elimina dicha alternativa.

Tabla 6. Ventajas y desventajas de las alternativas de solución propuestas

Alternativa	Ventajas	Desventajas
Estandarizar los términos de las direcciones para	Formato único que permite mejor entendimiento de las direcciones	Infinitas formas de los usuarios digitar las direcciones

crear un formato final único	Trazabilidad desde el formato único creado a la separación de componentes: Dirección y complemento	Diferentes formas de codificar las direcciones en el país
Realizar una macro que permita la transformación y separación de las direcciones	Modelamiento a la medida de la macro para que cumpla las necesidades actuales	Por ser un modelamiento manual se pueden pasar requerimientos por alto y quedar incompleta la macro
	Generar el input ideal para la georreferenciación en Google Maps	Infinitas formas de digitar la dirección lo cual puede influir en que no todo lo ingresado sea debidamente transformado
Crear un modelo de inteligencia artificial	Reconocimiento y transformación de datos instantánea creada por un robot	Cantidad de datos y tiempos extensos en entrenamiento al robot
	Porcentaje mínimo de error	Lenguaje de programación y herramientas de aplicación poco conocidas en la empresa.

Fuente. Elaboración propia.

4.3 Selección de la propuesta:

En la tercera etapa se desarrollaron diccionarios que permitieran la segmentación y estandarización de términos asociados a la dirección y al complemento, adicionalmente se creó una macro usando el lenguaje de expresiones regulares en VBA para la estructuración y separación automatizada de direcciones en gran volumen.

4.3.1 Propuesta para la estructuración de datos: inicialmente, se estableció que una dirección ideal para el proceso de entregas de última milla, está compuesta por una dirección estructurada y un complemento (ver **Tabla 3**), además se tiene claridad sobre componentes fundamentales que tiene cada una de estas partes, por lo que se decidió realizar un diccionario para cada una, en el cual se contengan las diversas formas de escritura asociadas a un formato estándar preestablecido.

Es importante segmentar y estructurar la dirección correctamente, dado que gracias a esto es que posteriormente se le podrá hacer la entrega oportuna al cliente de lo solicitado. La dirección estructurada es el insumo principal para la georreferenciación del punto de entrega, por esto es importante que el formato en que ingrese una dirección sea reconocido por la herramienta de georreferenciación, en este caso Google Maps. En la **Tabla 7**, se evidencia la estandarización de

términos para la dirección estructurada, cabe aclarar que se hizo una previa validación en Google Maps donde se verificó que los términos fueran leídos para obtener el punto destino.

Tabla 7. Diccionario de términos y abreviaturas para la dirección

Término estándar	Formas de escritura
CRA	Carrera, CARRERA, carrerra, carrerá, kra, kr, crra, karrera, etc
CL	CALLE, CLLE, Calle, calle, KL, KLLLE, KALLE, etc
TV	Transversal, TRANS, TRVS, TRASVERSAL, TR, trans, trs, etc
DG	diagonal, DIAGONAL, diag, diago, etc
#	NO, N°, Número, NUM, numero, no, núm, etc
AV	AVENIDA, ave, avenid, avda, etc
MZ	Manzana, Manz, manz, MANZANA, mza, maz
KM	kilómetro, kilometro, km, etc

Fuente. Elaboración propia.

A su vez, se realizó un diccionario de términos semejantes para el complemento de la dirección, ver **Tabla 8**, cuyo objetivo principal fue estandarizarlos y reconocer de qué forma puede ir compuesto el complemento para facilitar la separación de la dirección estructurada.

Tabla 8. Diccionario de términos y abreviaturas para el complemento

Término estándar	Formas de escritura
ED	Edif, edificio, ediff, edf, etc.
APTO	APARTAMENTO, ap, app, APART, apt, etc

URB	URBANIZACIÓN, urbanizacion, urba, urb., etc
INT	INTERIOR, inti, in, etc
PRIMER	1er, primero, etc
PISO	PS, piso, etc.
BG	BODEGA, bodeg, bodegas, BOD, etc
BR	BARRIO, brr, BAR, barrio, etc
...	...

Fuente. Elaboración propia.

Además, se deja la posibilidad de ir adicionando términos al diccionario, dado que hay n formas de digitarlas y cada que se reconozca una diferente puede ser agregada a cada diccionario con el propósito de ampliar la eficiencia de la propuesta planteada.

4.3.2 Propuesta para la transformación de datos: el proceso de transformación de términos y separación de dirección y complemento se desarrolló mediante la construcción de una macro en VBA de Excel, incluyendo los diccionarios anteriormente creados y el lenguaje de expresiones regulares, con el cual se procesa gran cantidad de datos en tiempos mínimos, obteniendo un resultado óptimo.

Inicialmente, para poder hacer uso de Expresiones Regulares en VBA Excel, se debe agregar la herramienta “Microsoft VBScript Regular Expressions 5.5”, luego de abrir el desarrollador, habilitarla mediante herramientas en VBA. Para aplicar esta metodología, se reconocieron las diferentes tipologías de las expresiones regulares, donde se agrupan determinados caracteres según la necesidad, ver algunos ejemplos en **Tabla 9**.

Tabla 9. Aplicación metodología Expresiones Regulares

Abreviatura	Contenido	Descripción
\d	[0-9]	Cualquier dígito
\D	[^0-9]	Cualquier cosa que no sea un dígito

\w	[a-zA-Z0-9_]	Coincide con cualquier letra, dígito o guión bajo
\W	[^a-zA-Z0-9_]	Cualquier cosa que no sea una letra, un dígito o un guión bajo.
\$		Coincide con el final de una cadena sin consumir ningún carácter.
*		Coincide con cero o más caracteres consecutivos.

Fuente. Elaboración propia.

Luego de reconocer los caracteres asociados a las diferentes cadenas de texto, se creó un código en VBA que automatizara la metodología manual de estructurar y separar la dirección del complemento, en este código se incluyeron los diccionarios para facilitar la transformación de términos semejantes y de este modo darle un formato estandarizado a la misma. Para esto, el código se dividió en # estructuras que se pueden evidenciar en la **Tabla 10**.

Tabla 10. Estructura del código en VBA.

Estructura	Descripción
Eliminar ruido	Parte del código se encarga de eliminar aquellos caracteres especiales que sólo generan ruido y no permiten el fácil reconocimiento de la información para su futura separación
Transformación de datos	Las siguientes líneas se encargan de transformar lo designado dentro de los diccionarios, es decir, reemplazando cada palabra contenida por su formato preestablecido.
Reconocimiento de partes	Esta parte del código asigna un carácter especial a cada palabra contenida en el diccionario de dirección estructurada y un carácter especial a los asociados al diccionario del complemento, con el fin de reconocer que hace parte de cada subdivisión de la dirección

Separación de partes	En estas líneas, el código se segmenta gracias a las expresiones regulares, donde se establece mediante caracteres como: “\d”, “\w”, “ * ” y aquellos asignados en el reconocimiento de partes, para segmentar que hace parte de la dirección y que hace parte del complemento.
Distribución en columnas	Finalmente, el código culmina enviando a una columna la dirección estructurada y la siguiente el complemento de la misma.

Fuente. Elaboración propia.

Una vez construido el código, consta sólo de ejecutarlo, para efectos prácticos se realizó una prueba piloto con la base de datos inicial de 450 direcciones, la cual fue ejecutada en 4.49 segundos, lo que reduce las horas-hombre en un 94,63%, sin embargo, hay direcciones que requieren intervención manual dado que el diccionario no reconoció la totalidad de palabras inmersas en las direcciones digitadas por los usuarios.

Para este caso en particular, la macro no procesó correctamente 24 direcciones, es decir, un margen de error del 5,33%, las cuales pueden ser procesadas manualmente en un tiempo de 721, 92 segundos y aun así sigue siendo eficiente su aplicación, dada la reducción de tiempo gracias a la automatización del proceso, se reduce el tiempo un 94,63%. Cabe aclarar que hay n formas de digitar las direcciones por lo que su automatización completa será complicada, pero entre más pruebas se apliquen se disminuirá el porcentaje de error dado que se agregan palabras al diccionario y de este modo se reconocen más direcciones.

Posteriormente, se deben georreferenciar las direcciones para lo cual se encontró una herramienta en Google Drive - Hojas de cálculo, llamada “Geocode by Awesome Table” la cual permite georreferenciar direcciones en tiempo cortos, para la base de datos de 450 direcciones ya bajo el modelo de dirección estructurada, tarde 319 segundos en encontrar tanto latitud y longitud por cada punto, sin embargo no logró localizar 8 puntos, es decir, tuvo un margen de error del 1,7%; sin embargo, georreferenciar estos punto de forma manual tarda 13954,5 segundos, es decir, el tiempo de ejecución disminuye 95,94%.

La automatización en el procesamiento de direcciones, permite la integración de diversas herramientas como la disminución de tiempos de ejecución, la estandarización global de términos asociados, además de enmarcar un nivel de error poco significativo el cual puede ser disminuido gracias a la posibilidad de seguir mejorando el modelo agregando diversas variables.

4.4 Comparación de métodos y toma de tiempos:

La cuarta etapa consistió en comparar el método actual con la propuesta planteada en aspectos como el tiempo de aplicación y la eficiencia de cada uno, con el propósito de determinar la funcionalidad y las variaciones en el procesamiento de direcciones para el sistema última milla. Para esto, se realizaron pruebas para recopilar datos y la aplicación de una encuesta a los integrantes del equipo de trabajo con el fin de evaluar la funcionalidad del método.

4.3.1 Comparación de métodos y toma de tiempos: al realizar un sondeo por el proceso inicial se evidenció que los puntos críticos se basaban en la baja automatización entre actividades, pues éstas estaban siendo realizadas de forma manual generando tiempos amplios de ejecución y generando una tarea engorrosa de realizar. A razón de esto, el proyecto se direccionó a mitigar la operatividad manual, automatizando actividades y aumentar la eficiencia operativa.

El método actual y el propuesto constan de la misma secuencia de actividades, la principal diferencia es que dos actividades: Reestructuración de direcciones y Georreferenciación de las mismas, se realizan de forma masiva en la propuesta planteada, mientras que actualmente se realizan de forma manual. Con la propuesta se logra disminuir un 82,59% el tiempo de ejecución del proceso.

Actualmente, para la Reestructuración de direcciones, se requiere la intervención una a una de cada dirección, donde se estandariza la información contenida; mientras que la propuesta contiene diccionarios que permiten la estandarización masiva de la información de las direcciones, donde procesar una base de datos de 450 direcciones se realiza en 726.41 segundos, (masivamente se logran procesar las direcciones en 4.49 segundos y las demás procesadas manualmente), lo que

reduce las horas-hombre en un 94,63% comparado con el método actual, esto se puede evidenciar en la **Tabla 11**.

Por otra parte, la Georreferenciación actualmente se realiza de forma manual, lo que genera tiempos elevados de ejecución, aproximadamente 31,01 segundos por cada punto, mientras que con la herramienta “Geocode by Awesome Table” se obtiene de forma masiva y en tiempo eficiente, donde se procesan 450 direcciones en 319 segundos, con un margen de error de 1,7%, es decir, 8 direcciones deben ser procesadas manualmente, lo que disminuye el tiempo de ejecución un 95,94%, ver **Tabla 11**.

Tabla 11. Comparación de tiempos entre métodos

Actividad	Comparación duración métodos (segundos) para una base de 450 datos		
	Método actual	Método planteado	Variación (%)
Limpieza de datos	13536	726.41	-94,63%
Georreferenciación	13954,5	567.08	-95,94%
Optimización:	1230	1230	0%
Estructuración de las hojas de ruta	1800	1800	0%
Seguimiento a las entregas	1200	1200	0%
Total	31720,5	5523,49	-82,59%

Fuente. Elaboración propia.

Con base en los resultados anteriores, se ve reflejada la intervención en los puntos críticos detectados en el método actual, lo que aumentó de forma significativa la eficiencia del proceso y generó valor para el procesamiento de direcciones en el sistema última milla. A su vez, se estructuró un procedimiento óptimo y de fácil intervención, lo que permite que cualquier integrante del equipo lo realice y actualice futuras mejoras.

4.3.2 Respuesta del equipo de trabajo: inicialmente, se citó una reunión con los integrantes del equipo para presentarles la propuesta planteada: la macro y la herramienta encontrada de Google Drive, además de explicarles la funcionalidad y las características específicas asociadas a la propuesta y de esta forma determinar el valor percibido mediante una encuesta basada en tres preguntas, enunciadas a continuación:

- ¿Cree usted que la propuesta planteada mejora el proceso actual?
- ¿Considera que lo propuesto es funcional y eficiente para el procesamiento de direcciones?
- ¿Considera que el proceso planteado es entendible y práctico?
- Califique su satisfacción con la propuesta planteada, donde 1 es insatisfecho y 5 es muy satisfecho.

Según los resultados obtenidos a partir de la encuesta, se encontró que el 100% de los integrantes del equipo consideran que la propuesta planteada mejora el proceso actual de estructuración y georreferenciación de direcciones en el sistema de última milla, como se puede observar en la **Figura 2**. Entre los principales comentarios alusivos a la propuesta se destacan la eficiencia, la claridad y el ahorro de tiempo que proporciona esta alternativa.

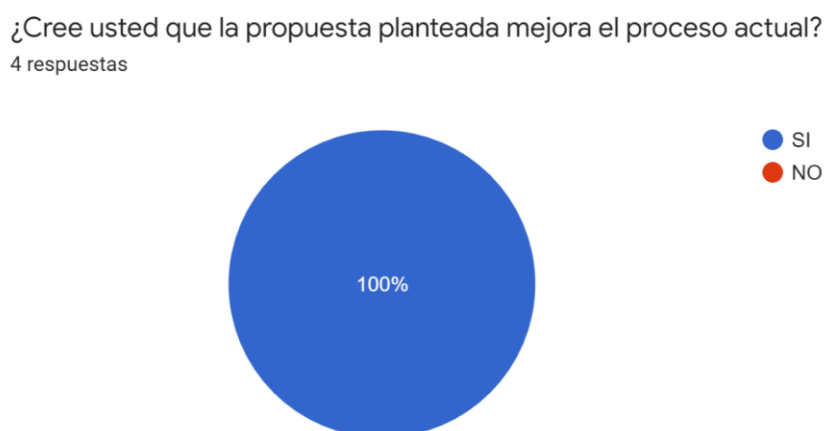


Figura 2. *Mejoramiento del proceso actual.*

Nota. Fuente elaboración propia.

Además, el 100% del equipo consideró que la propuesta planteada es eficiente y funcional para el procesamiento de direcciones, además que es entendible y práctica, esto se puede observar en la **Figura 3** y **Figura 4**, entre los comentarios destacados se tiene la fácil aplicabilidad de la propuesta y la posibilidad de replicar el modelo en otras zonas con sólo la variabilidad de modificar los diccionarios.

¿Considera que lo propuesto es funcional y eficiente para el procesamiento de direcciones?
4 respuestas

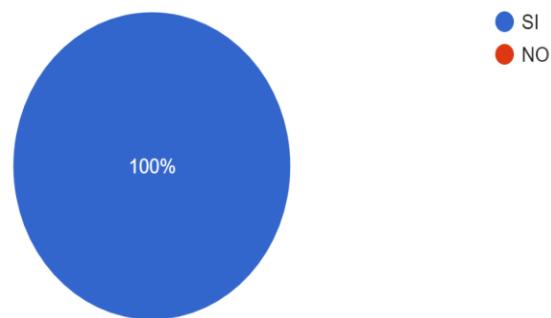


Figura 3. *Funcionalidad y eficiencia del proceso*

Nota. Fuente elaboración propia.

¿Considera que el proceso planteado es entendible y práctico?
4 respuestas

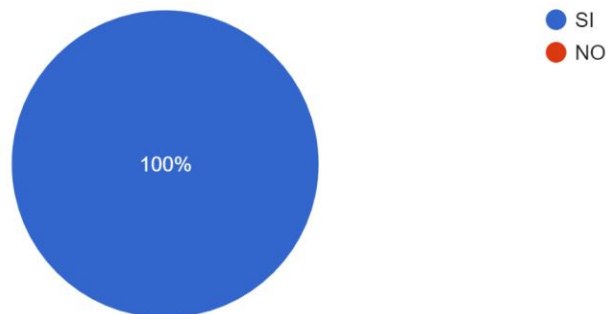


Figura 4. Entendimiento del proceso

Nota. Fuente elaboración propia.

Por último, en la **Figura 5** se evidencia que el 75% del equipo calificó la propuesta en 5 la satisfacción de la propuesta, en una escala de 1 a 5, mientras que el 25% la calificó en 4. Este resultado es muy alentador, dado que fue bien recibida la propuesta por el equipo, además reconocen la aplicabilidad que tendría de modo transversal en la compañía, la reducción de costos y tiempos de ejecución de la actividad, además del valor agregado que le brinda al procesamiento de direcciones.

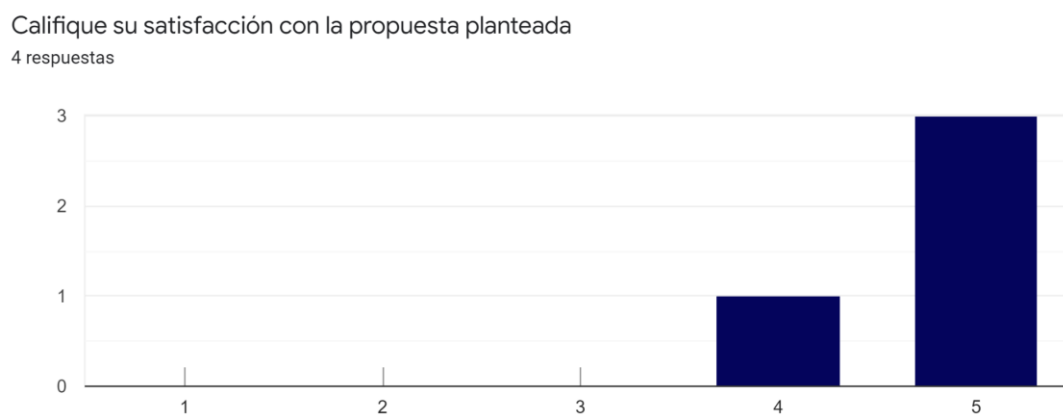


Figura 5. Satisfacción de la propuesta

Nota. Fuente elaboración propia.

Los resultados obtenidos con la implementación de la propuesta que permite automatizar el procesamiento de direcciones tanto en la normalización de cada dirección como en la georreferenciación de las mismas, fueron la disminución los tiempos de ejecución en un 82,59% con un margen de error inferior al 6%, lo cual fue bien recibido por el equipo de trabajo, quienes avalaron la propuesta y resaltaron el valor agregado que le dio al proceso.

5. Conclusiones

El COVID-19 no sólo obligó al mundo a resguardarse, sino que obligó a las empresas a transformarse, pues se pasó de comprar directamente en las tiendas, a la comodidad de recibir los productos en las casas, mediante un sistema de entrega denominado última milla, la cual asegura que el 70% de los compradores repiten una compra dada la satisfacción que hayan tenido con la recepción de su pedido. El Grupo Familia, una compañía líder en higiene y salud a nivel global, se sumó a la transformación y modificó su cadena de suministro para llegar al consumidor final, en la cual identificó que una de las principales problemáticas radica en la información suministrada por los clientes, específicamente las direcciones destino.

Con base en el análisis realizado, se encontró que las diferentes marcas del Grupo Familia captan la información de sus clientes mediante formularios de inscripción de las páginas web, cuyos campos de llenado son limitados, lo que implica que la información quede agrupada. El proceso actual de transformación de datos para la posterior entrega mediante el sistema de última milla consta de cinco fases fundamentales, de ellas la fase limpieza de datos tarda alrededor de 30.08 segundos y la fase georreferenciación 31.01 segundos, ya que no están automatizadas y se realizan de manera manual incrementando el tiempo de operación al sumarse los tiempos del total de usuarios registrados.

De acuerdo con lo anterior, se agruparon aquellas ideas de automatización que permitieran dar solución a las problemáticas anteriormente mencionadas. Se encontraron alternativas como el cambio de los campos de los formularios, modelar la información recolectada, ya sea mediante una macro en Excel o mediante el uso de Inteligencia Artificial. La implementación de una macro de Excel con lenguaje de Expresiones Regulares fue la alternativa más práctica y eficiente, ya que permite la transformación masiva de información, además fue posible su realización en el tiempo pactado y fácil de replicar por el equipo de trabajo.

Con el fin de darle solución a la problemática de los datos, se desarrolló un código en Excel basado en Expresiones Regulares, que se encarga de eliminar aspectos innecesarios de la información, segmentar y estandarizar palabras claves de la dirección y el complemento para llevar

a cabo la separación de estas. Se desarrolló una alternativa de automatización para el procesamiento de direcciones que permite la transformación del campo de dirección digitada por los usuarios en pocos segundos, procesando una base de datos de 450 usuarios en 4.49 segundos con un margen de error del 5.33%, generando dos únicos datos: uno asociado a la dirección estructurada (CL 123 #43-28), que sería el input para la georreferenciación y otro, el complemento (Apto 501, Ed El portal), que permite al transportador ubicar en el punto al consumidor final. Estos datos transformados pueden ser utilizados en Geocode, una herramienta de Google Drive que georreferencia direcciones de forma masiva.

Con la implementación de esta alternativa de automatización en una base de datos de 450 usuarios, se encontró una disminución del 82,59% en el tiempo de ejecución del proceso. Para la limpieza de datos se obtuvo un tiempo operativo de 726,41 segundos comparado con el tiempo asociado del método inicial, que era de 13536 segundos, alcanzando una eficiencia del 94,67%. A su vez, el proceso de georreferenciación pasó de requerir 13954,5 segundos a 567,08 segundos con una eficiencia del 98,3%. Cifras muy alentadoras para la aplicación de la propuesta, lo que es satisfactorio para el equipo de trabajo, quienes aprueban y valoran la utilidad que tiene esta propuesta en el procesamiento de direcciones en el Grupo Familia.

Según los resultados, se logró mejorar a nivel general el procesamiento de direcciones para el sistema de última milla, ya que se logró automatizar las actividades que implicaban una completa operatividad manual y que fueron identificados inicialmente como los problemas principales. Con respecto a la aplicabilidad y replicabilidad de la propuesta, es apta para cualquier integrante del equipo y es posible su replicabilidad en otras zonas del mundo, además de que disminuye los costos asociados a la mano de obra y la estandarización del proceso.

Referencias

- Belinchón Monjas, Y. (2019). *Minería de datos*. <https://www.it.uc3m.es/jvillena/irc/practicas/10-11/15mem.pdf>
- Chapman, A. D. (2005). Primary species and species-occurrence data: Principles and methods of Data Cleaning.
- Cortés, P. (2021, January 11). La importancia de la última milla en tiempos de pandemia. <https://www.revistalogistec.com/vision-empresarial/analisis-2/3175-la-importancia-de-la-ultima-milla-en-tiempos-de-pandemia>
- Dávila Martínez, F. J., & Camacho Arranz, E. (2012, 5 octubre). Georreferenciación de documentos cartográficos para la gestión de Archivos y Cartotecas. Vibercarto. <https://www.ign.es/web/resources/docs/IGNCnig/CTC-Ibercarto-V-Georreferenciacion.pdf>
- Dib, F. (s. f.). *Regex101: build, test, and debug regex*. Regex101. Recuperado 28 de febrero de 2022, de <https://regex101.com/>
- González, M. (2019). Empresas y proveedores corren para ganar la «última milla». Recuperado el 28 de octubre de 2021 de: http://www.unionprofesional.com/clipping/131219/LaRazon_081219.pdf
- Humberto, F., & Rozas, N. E. F. (2021). Limpieza, corrección y geocodificación de grandes bases de direcciones utilizando minería de texto. *Universidad Ciencia y Tecnología*, 25(109), 80-87.
- Importancia de normalización de direcciones en BBDD*. (2021, 25 noviembre). DEYDE. Recuperado 19 de marzo de 2022, de <https://deyde.com/news/normalizacion-direcciones-base-datos/>
- Lee, M. L., Hsu, W., & Kothari, V. (2004). Cleaning the Spurious Links in Data. *IEEE Intelligent Systems*, 19(2), 28–33. <https://doi.org/10.1109/MIS.2004.1274908>
- Martínez Izquierdo, E. (2015). *Inteligencia artificial aplicada al análisis de datos*. https://oa.upm.es/40381/1/INVE_MEM_2015_209053.pdf
- Ocampo D. (2021). Entregable del objetivo reconocimiento de actividades para llevar a cabo el proyecto de “Implementación de herramientas de SIG y transformación de datos para disminuir los

tiempos de ejecución de las actividades de data cleaning, sectorización y seguimiento en el proceso de entregas a cliente final del área de negocios logísticos de la empresa Grupo Familia S.A. ubicada en la ciudad de Medellín”.

Overview / Geocoding API /. (s. f.). Google Developers. Recuperado 19 de marzo de 2022, de <https://developers.google.com/maps/documentation/geocoding/overview>

Somos Grupo Familia. (n.d.). Retrieved October 28, 2021, from <https://www.grupofamilia.com/somos-grupo-familia>

Regex — Cómo usar expresiones regulares (Regex) en Microsoft Excel tanto en la celda como en loops. (2014, 20 marzo). regex. Recuperado 28 de febrero de 2022, de <https://www.it-swarm-es.com/es/regex/como-usar-expresiones-regulares-regex-en-microsoft-excel-tanto-en-la-celda-como-en-loops/1044621562/>

Anexos

Anexo 1, Diagrama de flujo del proceso actual

