

Práctica empresarial analista de base de datos

Javier Andrés Gutiérrez Araujo

Informe de Practica para optar el título de Ingeniero Mecatrónico

Director

Carlos Daniel Pimentel Díaz

Maestria en Inteligia Artificial

Universidad Santo Tomás, Bucaramanga

División De Ingenierías Y Arquitectura

Facultad De Ingeniería Mecatrónica

2024

Contenido

Introducción	7
1 Práctica empresarial analista de base de datos.....	9
1.1 Reseñas de la Empresa.....	9
1.2 Misión	9
1.3 Visión.....	10
2 Justificación	10
3 Identificación de la problemática.....	10
4 Solución de la problemática.....	11
5 Marco de referencia	17
5.2 Gestión de datos.....	17
5.3 Relevancia de la base de datos.....	18
6 Objetivos	19
6.1 Objetivo general.....	19
6.2 Objetivos específicos	19
7 Plan de trabajo.....	19
Referencias.....	22

Lista de figuras

Figura 1. Mensaje actualización planta de la campaña adultos.....	11
Figura 2. Mensaje de actualización planta campaña México.....	12
Figura 3. Proceso antiguo para actualizar las plantas.....	13
Figura 4. Proceso para actualizar las plantas mediante Python.....	14
Figura 5. Indicador AHT para la campaña de Latam.	15
Figura 6. Indicador AHT para la campaña de Brasil.	15
Figura 7. Mensaje para Asesores	16
Figura 8. Mensaje para Team Leader.	16
Figura 9. Mensaje para Team Leader Brasil.	16
Figura 10. Mensaje para Team Leader Latam.....	17
Figura 11. Esquemático de actividades.	21

Resumen

Este informe detalla las actividades realizadas durante la práctica empresarial en Open English, centradas en la mejora de la eficiencia operativa a través de la implementación de soluciones tecnológicas. Se abordaron desafíos relacionados con la gestión manual de datos y procesos, como la actualización de plantillas y la asignación de auditorías, utilizando herramientas como Python para desarrollar automatizaciones que optimizaron estas tareas. El resultado fue una notable reducción en los tiempos de procesamiento y una mayor exactitud en la administración de datos. Además, se monitorearon y analizaron indicadores clave de desempeño en campañas de Latam y Brasil, contribuyendo a una mejor toma de decisiones y al incremento de la productividad organizacional.

Palabras clave: eficiencia operativa, automatización, análisis de datos, Open English, mejora de procesos

Abstract

This report outlines the activities carried out during the business internship at Open English, focusing on enhancing operational efficiency through the implementation of technological solutions. Challenges related to manual data management and processes, such as updating templates and assigning audits, were addressed by developing Python-based automations that optimized these tasks. The outcome was a significant reduction in processing times and improved data management accuracy. Furthermore, key performance indicators in Latam and Brazil campaigns were monitored and analyzed, contributing to better decision-making and increased organizational productivity.

Keywords: operational efficiency, automation, data analysis, Open English, process improvement

Glosario

Automatización: proceso de usar tecnología para realizar tareas con mínima intervención humana, aumentando la eficiencia y reduciendo errores.

Análisis de datos: proceso de inspeccionar, limpiar y modelar datos con el objetivo de descubrir información útil, sacar conclusiones y apoyar la toma de decisiones.

API: conjunto de definiciones y protocolos que permite que un software se comuniquen con otro.

Google Sheets: aplicación de hoja de cálculo en línea ofrecida por Google, que permite la colaboración y el trabajo en tiempo real.

Python: lenguaje de programación ampliamente utilizado en el análisis de datos, desarrollo web, automatización de tareas, entre otros.

Indicador de Productividad: métrica utilizada para evaluar la eficiencia de los empleados en el cumplimiento de sus tareas, especialmente en entornos de campaña.

Raspberry Pi: pequeño ordenador de bajo costo que se utiliza para proyectos de programación, automatización y robótica.

Visual Studio Code: editor de código fuente que soporta múltiples lenguajes de programación y es utilizado comúnmente para el desarrollo de software.

Introducción

La presente investigación tiene como objetivo principal analizar y optimizar los procesos de Open English, una plataforma líder en educación en línea, mediante la implementación de herramientas tecnológicas para eliminar las barreras geográficas y de horario. En un entorno donde la gestión eficiente de datos es esencial para el éxito organizacional, se hace evidente la necesidad de mejorar la precisión, velocidad y eficiencia en el manejo de la información.

En este contexto, se identifica la falta de automatización en ciertos procesos clave, como la actualización de plantillas de personal, asignación de dirigidos y auditorías, y conciliación de auditorías presentadas. Esta situación ha generado desafíos significativos, incluyendo la propensión a errores, baja eficiencia en procesos manuales y limitada capacidad para manejar grandes volúmenes de datos de manera efectiva.

Ante esto, se implementó el lenguaje de programación Python para la automatización, como una solución necesaria para optimizar los procesos de análisis de datos. En este contexto, la automatización consiste en tomar datos de una hoja de Google Sheets a través de la API, procesarlos en Python y luego subir la información de nuevo a Google Sheets o generar una alerta mediante la aplicación de Slack. Este informe aborda la identificación de la problemática, la justificación de la necesidad de automatización, y los objetivos específicos para mejorar la eficiencia operativa y la toma de decisiones dentro de la organización.

Se logró optimizar procesos como la actualización de las plantas por campaña, la consolidación de datos y el traspaso de la plataforma de Microsoft Access a Google Sheets. Además, se implementó un seguimiento a un indicador de productividad en las campañas de Latam y Brasil, con notificaciones personalizadas. La actualización de las plantas se redujo de una hora a solo 5 u 8 minutos. En cuanto a la consolidación y traspaso de información de Access a Google

Sheets, el tiempo de ejecución se redujo a segundos, comparado con los 20 a 30 minutos que tomaba anteriormente.

1 Práctica empresarial analista de base de datos

1.1 Reseñas de la Empresa

Open English es una plataforma líder en educación en línea que ofrece cursos de idiomas flexibles y personalizados para estudiantes de todos los niveles. Con su enfoque innovador, la plataforma combina clases en vivo impartidas por profesores nativos de habla inglesa con contenido interactivo y herramientas tecnológicas. Esto permite a los estudiantes aprender el idioma que desean desde la comodidad de sus hogares o cualquier lugar con conexión a internet, eliminando las barreras geográficas y de horario.

La plataforma ofrece una amplia variedad de recursos, que incluyen lecciones en video, ejercicios de gramática, práctica de conversación y evaluaciones de progreso. Además, los estudiantes tienen acceso a un tablero personalizado donde pueden dar seguimiento a su progreso, programar clases y recibir retroalimentación de los profesores.

La flexibilidad es otra ventaja clave de Open English, ya que los estudiantes pueden elegir entre una variedad de planes de estudio y horarios de clases que se adaptan a sus necesidades y horarios. Además, la plataforma ofrece soporte técnico y académico las 24 horas del día, los 7 días de la semana, para garantizar una experiencia de aprendizaje sin contratiempos [4].

1.2 Misión

Transformando el aprendizaje del inglés

La Misión de Open English es reinventar la manera de aprender inglés, derribando las barreras que impone la enseñanza tradicional [3].

1.3 Visión

Cambiando la vida de la gente

Nuestra visión es cambiar vidas a través del poder del idioma global [3].

2 Justificación

La necesidad de automatización en los procesos de análisis de datos se fundamenta en la creciente complejidad y volumen de la información que las empresas deben gestionar en la actualidad. La falta de automatización conlleva una serie de desafíos, incluyendo la propensión a errores humanos, la baja eficiencia en los procesos manuales y la limitada capacidad para manejar grandes volúmenes de datos de manera efectiva. La implementación de herramientas como Excel y Google Sheets, si bien son útiles en ciertos aspectos, resultan insuficientes para abordar las demandas actuales de análisis de datos debido a sus limitaciones en términos de escalabilidad, flexibilidad y capacidades de procesamiento.

Por lo tanto, implementar un lenguaje de programación como Python se presenta como una solución necesaria para optimizar los procesos de análisis de datos, permitiendo una mayor velocidad, precisión y eficiencia en la manipulación y procesamiento de la información.

3 Identificación de la problemática

La falta de automatización de procesos relevantes como la actualización de las plantas del personal que estaría trabajando en cada campaña que tiene la empresa, la asignación de auditorías y dirigidos, la conciliación de las auditorías para el cálculo de las comisiones de ventas y servicios; y la modificación de errores, conlleva a que todo este tipo de actividades se realicen de forma manual, copiando y pegando la información de un archivo de Excel a una hoja de cálculo en Google Sheets.

Para los procesos que implica realizar el cálculo de las comisiones para los asesores, estos se realizan directamente en las mismas hojas de Google Sheets haciendo que los archivos sean bastante lentos cuando se tiene mucha información cargada en dichas hojas de cálculo.

Las principales razones por la que algunas empresas manejan sus bases de datos, consolidados y realizan cálculos directamente en hojas de cálculo como Excel o Google Sheets es por la facilidad de uso, la amplia disponibilidad ya que en la gran mayoría de computadores y portátiles ya viene preinstalado Excel, adicional a esto la curva de aprendizaje es menos demandante para los analistas en comparación con Python, Se considerarían ventajas, pero en tema de rendimiento y automatizar dichas aplicaciones mencionadas son insuficientes. Otra desventaja que presentan es la escalabilidad, cálculos completos y manejo de volúmenes grandes de datos, Excel y Google Sheets no son eficientes para dichas tareas y pueden tardar varios minutos o incluso horas. Todas estas desventajas afectan la productividad y la eficiencia del equipo de trabajo.

4 Solución de la problemática

Inicialmente, se propuso implementar un servidor MySQL en una Raspberry Pi 4. Sin embargo, por razones de seguridad, no fue posible conectar las máquinas de la empresa con el servidor en la Raspberry Pi 4, incluso estando en la misma red.

En su lugar, se optó por instalar Visual Studio Code en el PC proporcionado por la empresa para usar Python en la automatización de los procesos mencionados en esta investigación. Para conectar Google Sheets con Python, se utilizó la API oficial de Google tanto para extraer como para subir datos. Para interactuar con Excel, se emplearon las librerías oficiales de Python, permitiendo obtener datos y crear nuevos archivos de Excel. Además, se usó la API oficial de Slack para enviar mensajes de tipo alerta a canales específicos o a usuarios concretos.

En las figuras 1 y 2 se puede ver como se presenta un mensaje sobre la actualización de las plantas, si en las campañas pertinentes se generó algún retiro o, por el contrario, no se encontró alguna novedad.

Figura 1. *Mensaje actualización planta de la campaña adultos*

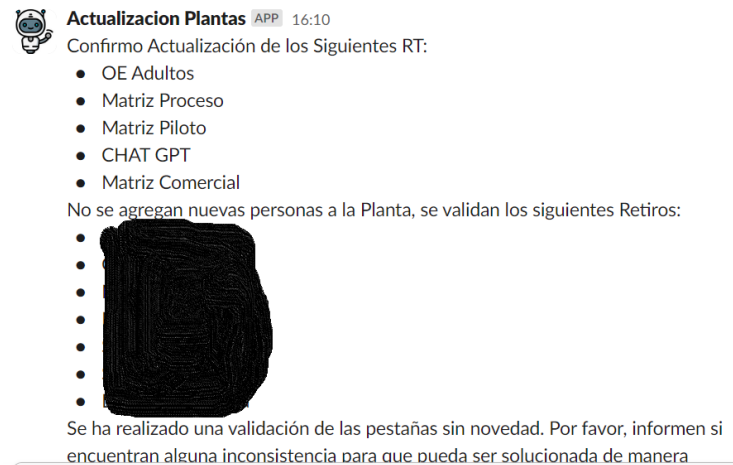


Figura 2. *Mensaje de actualización planta campaña México*

Confirmo Actualización de los Siguietes RT:

- OE Adultos Mexico

No se encuentran Novedades, ni personas para ingreso en campaña.

Se ha realizado una validación de las pestañas sin novedad. Por favor, informen si encuentran alguna inconsistencia para que pueda ser solucionada de manera oportuna.

Tener en cuenta que la actualización se realiza con HC actualizado el día de hoy

Por favor, estén atentos a cualquier actualización adicional y no duden en comunicarse si tienen alguna pregunta o inquietud.

La empresa también tiene bases de datos alojadas en servicios como Five9 y Salesforce, que cuentan con APIs para conectar con Python. Estas APIs permiten obtener datos brutos directamente en Python y realizar el procesamiento necesario. Luego, los datos pueden subirse a archivos de Google Sheets y generar mensajes de tipo alerta para los líderes de equipo de las respectivas campañas sobre el rendimiento de las operaciones.

En la figura 3 se observa cómo se realizaba el proceso para actualizar las plantas de las diferentes campañas que maneja la compañía, este proceso se hacía de manera manual y se debía realizar por cada planta que se va a actualizar, excepto la descarga del archivo HC máster que solo se hace una vez.

En la figura 4 se observa el proceso realizado por Python para actualizar las plantas de cada campaña, el cual no tiene interacción humana durante este proceso, ya que se cuenta con una tarea programada en el portátil que se ejecuta todos los lunes.

Figura 3. Proceso antiguo para actualizar las plantas.

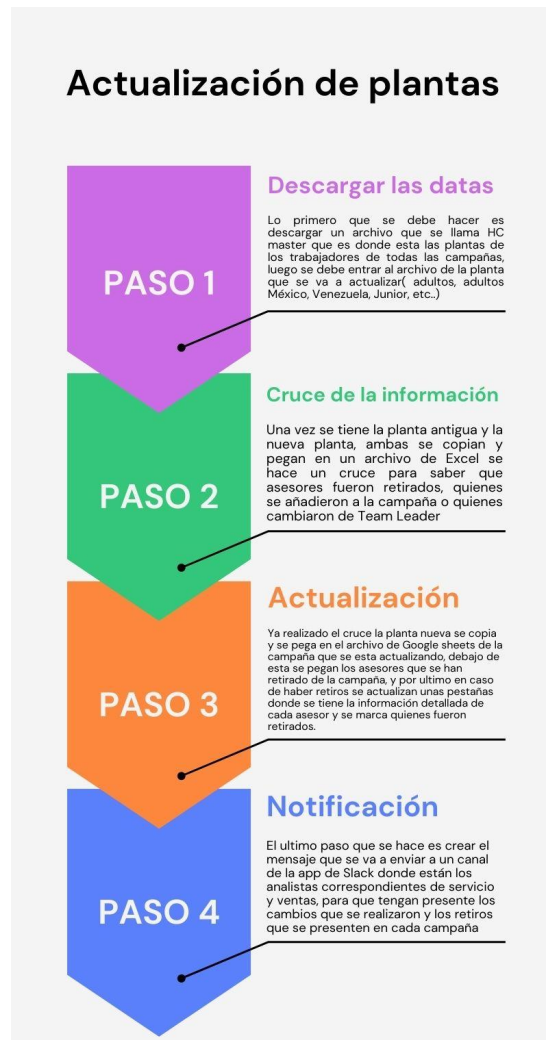
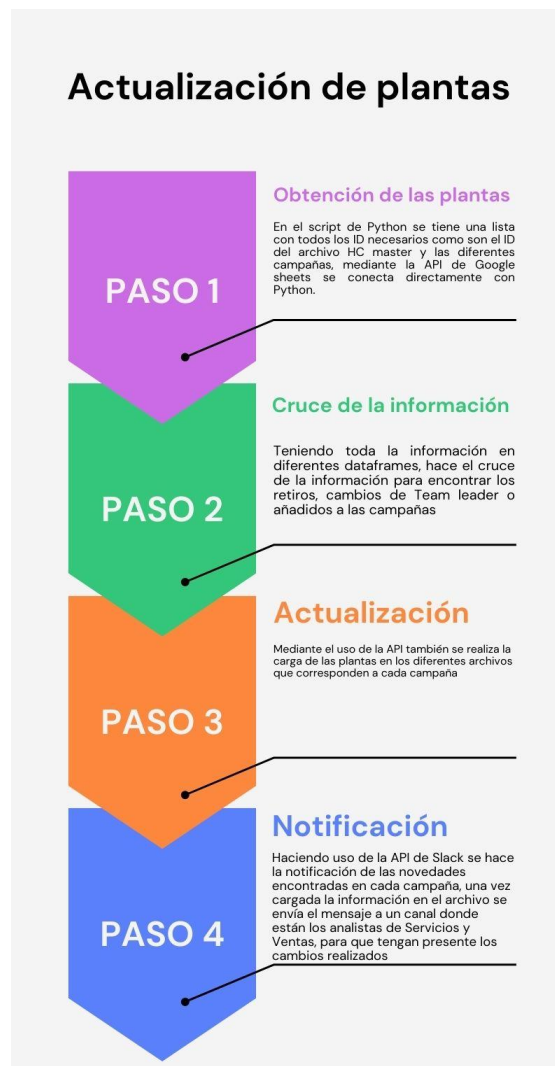


Figura 4. *Proceso para actualizar las plantas mediante Python.*

El proceso de manera manual como se hacía antes tardaba 1 hora para actualizar todas las plantas y generar la notificación, con la implementación de Python en este proceso se logró reducir el tiempo alrededor de 5 o 8 minutos.

Adicional a lo mencionado anteriormente gracias a mis practicas se llevó a cabo un proyecto que consiste en darle seguimiento a un indicador de productividad de los asesores de la compañía en las campañas de Latam y Brasil, dicho indicador es recopilado hora a hora todos los días por un gestor de base de datos llamado Five9, en dicho proyecto se

obtiene la información en bruto, se procesa, se limpia y se realizan los cálculos pertinentes que funcionan como activador de la alerta.

Se puede observar en la figura 5 el indicador ya procesado para la campaña de Latam y en la figura 6 el indicador ya procesado para la campaña de Brasil.

Figura 5. *Indicador AHT para la campaña de Latam.*

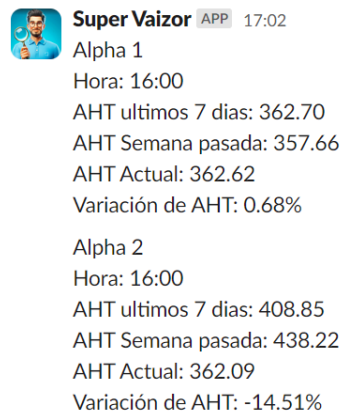


Figura 6. *Indicador AHT para la campaña de Brasil.*

Alpha BR 1
Hora: 07:00
AHT ultimos 7 dias: 369.89
AHT Semana pasada: 404.69
AHT Actual: 243.22
Variación de AHT: -37.20%

Alpha BR 2
Hora: 07:00
AHT ultimos 7 dias: 444.41
AHT Semana pasada: 429.20
AHT Actual: 477.35
Variación de AHT: 9.28%

Si la campaña genera una alerta que se estableció para valores donde la Variación de AHT sea menor al 4% se envían mensajes para los asesores de manera interna (Figura 7) y también para los Team Leader. (Figuras 8, 9 y 10)

Figura 7. Mensaje para Asesores

Hola [REDACTED] ⚠️
En el día has recibido un total de 17 llamadas, y tu AHT es 170 Seg. Estas -40.2% por debajo del AHT de tus compañeros del mismo Tier en la hora pasada. Deberías tener un AHT (Llamadas de 30 seg) acumulado de 335 seg

Figura 8. Mensaje para Team Leader.

AHT Alert
🎉👏
Hola [REDACTED], estas haciendo un gran trabajo logrando que tus asesores extiendan de manera optima

Figura 9. Mensaje para Team Leader Brasil.

AHT Alert
OE Brazil 🙄
Hora de la Desviación: 10:00
Hola [REDACTED], lo estas haciendo bien. Sin embargo, tienes oportunidad de mejora en la extensión de llamadas mayores a 30 seg con estos asesores ya que son los que tienen el AHT más bajo en sus respectivos tiers

AGENT NAME	Tier	Gerente	AHT	conectadas
[REDACTED]	Alpha 3 BR	[REDACTED]	2:50	17
[REDACTED]	Alpha 3 BR	[REDACTED]	3:29	15

Figura 10. Mensaje para Team Leader Latam.

AHT Alert
 OE Latam ⚠
 Hora de la Desviación: 08:00
 Hola [REDACTED], Estos asesores no estan extendiendo lo suficiente las llamadas alpha en comparacion al resto de sus compañeros del mismo tier

AGENT NAME	Tier	Gerente	AHT	conectadas
[REDACTED]	Alpha 2	[REDACTED]	5:52	12
[REDACTED]	Alpha 2	[REDACTED]	6:44	11
[REDACTED]	Alpha 2	[REDACTED]	5:14	8
[REDACTED]	Alpha 2	[REDACTED]	6:49	8
[REDACTED]	Alpha 3	[REDACTED]	6:15	5
[REDACTED]	Alpha 3	[REDACTED]	5:48	4

Como se puede apreciar en las imágenes anteriores los diferentes mensajes que se configuraron para los Team Leader, el formato de colores es igual para ambas campañas y este depende de cuantos asesores están generando la alerta, Verde es que no tienen ningún asesor afectado, naranja para los que tienen menos de 3 asesores y rojo para los Team Leader que tienen más de 3 asesores.

5 Marco de referencia

5.2 Gestión de datos

La gestión de datos se refiere a la organización, almacenamiento y procesamiento de información de manera eficiente y efectiva. En el contexto de este estudio, la gestión de datos es crucial para garantizar la integridad, disponibilidad y confidencialidad de la información recolectada. Se explorarán diversas estrategias y mejores prácticas en la gestión de datos para optimizar su utilización y análisis [5].

Beneficios

- Ventaja competitiva
- Decisiones más rápidas y acertadas.
- Innovación y reducción de riesgos.
- Aumento de la productividad.

5.3 Relevancia de la base de datos

La base de datos seleccionada para este estudio es un componente crítico en la recopilación y análisis de datos. Se espera que la base de datos proporcionada permita almacenar grandes volúmenes de datos de manera eficiente y escalable. Además, se explorarán las funcionalidades y características específicas de la base de datos en relación con los objetivos de investigación, como la capacidad de consultas complejas, la seguridad de los datos y la capacidad de integración con otras herramientas de análisis [1].

Funcionalidades de la gestión de datos

- Automatización de procesos.
- Nube multifuncional.
- Observabilidad de los datos

6 Objetivos

6.1 Objetivo general

Mejorar la eficiencia operativa por medio de herramientas de análisis y visualización de datos para generar insights accionables, identificar oportunidades de mejora y aumentar la productividad de la empresa Open English.

6.2 Objetivos específicos

- Actualizar las plantas del personal para cada campaña mediante el lenguaje de programación Python, asegurando que la información sea precisa y actualizada, para lograr una asignación eficiente de recursos y una gestión efectiva del personal en todas las campañas que la empresa lleva a cabo.
- Optimizar la asignación de dirigidos y auditorías dentro de la empresa, utilizando la herramienta Python con el propósito de mejorar la eficiencia operativa.
- Realizar la conciliación de las auditorías presentadas, con la ayuda del lenguaje de programación Python con el fin de identificar y resolver discrepancias de manera oportuna, asegurando su exactitud y transparencia.

7 Plan de trabajo

Las actividades establecidas durante la práctica para dar cumplimiento a los objetivos específicos se realizaron de forma periódica, por lo que tuvieron una fecha de inicio y una fecha final cuando acabó mi periodo de prácticas. Dichas actividades se realizaron de manera simultánea desde el primer día que iniciaron mis prácticas.

Actividades realizadas

- Leer las bases de datos desde las hojas de cálculo Google Sheets mediante Python para ser procesados y segmentados en información útil para las diversas campañas.
- Exportar la información ya procesada a los archivos compartidos con los responsables de las áreas de calidad.
- Corrección de errores: para mantener las bases de datos actualizadas, se deben corregir ciertos errores humanos que cometen los analistas al momento de subir la información a las bases, una vez se escale y se aprueban los cambios de deben realizar y de esta manera estar al día con la información.
- Implementar el lenguaje de programación Python para automatizar los procesos de actualización de plantas, asignaciones (dirigidos y auditorías), consolidados de auditorías, esto se logró con la herramienta de Visual Studio Code, y extensiones de Jupyter, Python, extensiones externas como Google, Google OAuth para verificación de cuentas.

Las actividades necesarias para cumplir cada objetivo específico, como son actividades periódicas algunas semanales y otras diarias se hace entregable a mi supervisor que valida la realización de mis funciones. En la figura 11 se pueden apreciar los objetivos a cumplir y las actividades necesarias para alcanzarlos.

Figura 11. Esquemático de actividades.

Objetivos específicos	Evaluación de objetivos	Actividades realizadas para consecución de dichos objetivos	Fecha Inicial	Fecha Final	Entregable
1. Actualizar las plantas del personal en las diferentes campañas que maneja la empresa	Se identifica la planta actual dentro de la base de datos de la campaña a actualizar.	Realizar la actualización de las base de datos	19/02/2024		Informe Semanal
	Se descarga la base de datos que contiene los activos de todas las campañas de la empresa.				
	Se realizan cruces de ambas base de datos mencionadas para saber que personal ha sido contratado y no.				
	Se sube la base de datos ya actualizada.			18/08/2024	
2. Asignacion de base de datos(Dirigidos y auditorias)	Se identifica la base de datos donde estan se lleva un registro de todos los dirigidos y auditorias.	Realizar las asignaciones de los dirigidos y auditorias	19/02/2024		Informe diario
	Se identifican las diferentes bases de datos donde se sube la informacion				
	Se sube la informacion actualizada correspondiente a cada campaña y cada auditoria.			18/08/2024	
3. conciliacion de auditorias presentadas dentro de la compañía y de Comisiones del area de ventas y servicio	Se realizan reuniones en las que se identifican los errores	Actualizacion de informacion inconsistente	19/02/2024		Informe diario
	Realizar validacion de los monitores realizados en el	Verificar la informacion necesaria para obtener la calificacion de los asesores			Informe Mensual
	Mantener las bases de datos actualizadas			18/08/2024	
4.Modificacion de errores dentro de las e	Mediante unos formularios que estan vinculados a unas bases de datos especificas se realizan cambios en los errores	Modificacion de los errores	19/02/2024		Informe diario
	Para el mantenimiento de las bases de datos mensualmente se realizan cambios en los rangos que sacan la	Verificon de la informacion en las bases de datos		18/08/2024	Informe Mensual

Referencias

1. *Beneficios de la gestión de base de datos y análisis en un negocio* | SEIDOR. [En línea]. Disponible en: <https://www.seidor.com/blog/beneficios-gestion-base-datos-analisis-negocio>.
2. *Beneficios y funcionalidades clave de la gestión de datos este 2023*. [En línea]. Disponible en: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/beneficios-y-funcionalidades-clave-de-la-gestion-de-datos-este-2023>.
3. *Nuestros Valores - La importancia de Vision y Valores* | Open English. [En línea]. Disponible en: <https://www.openenglish.com/corporativo/mision-y-valores/>.
4. *Open English: opiniones, comentarios y reseñas (2024)*. [En línea]. Disponible en: <https://www.lapatria.com/educacion/open-english-opiniones>.
5. *¿Qué es la gestión de datos?* | Oracle Colombia. [En línea]. Disponible en: <https://www.oracle.com/co/database/what-is-data-management/>.