



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

ANÁLISIS DE DATOS PARA LA AUTOMATIZACIÓN DE
PROCESOS Y OBTENCIÓN DE INFORMES QUE
AYUDEN A LA TOMA DE DECISIONES.

INFORME DE ACTIVIDADES PROFESIONALES

QUE PARA OBTENER EL TÍTULO DE:

INGENIERO EN COMPUTACIÓN

PRESENTA

GARCÍA MARTÍNEZ VICENTE

ASESOR DE INFORME

DR. IVÁN VLADIMIR MEZA RUIZ



Ciudad Universitaria, Cd. Mx., 2022

Agradecimientos

Agradezco a mi familia, profesores, amigos y compañeros que estuvieron a lo largo de mi formación tanto académica como profesional ya que gracias a su paciencia y confianza logré terminar mis estudios y poder aplicarlos en la vida laboral tratando de mejorar los procesos actuales.

Índice General

Índice General	2
Resumen	3
1. Introducción	4
1.1 Objetivos	4
1.2 Estructura del informe	5
2. Análisis de Datos y herramientas	6
2.1 CRISP-DM	8
2.2 Análisis de datos en las Empresas	10
2.3 Data Studio	11
2.4 Google Apps Script	12
2.5 Desarrollos en la empresa	13
3. Automatización de proceso de notificaciones	14
3.1 Entendimiento de requerimientos	15
3.2 Comprensión y obtención de los datos	16
3.3 Preprocesamiento de los datos	17
3.4 Análisis	18
3.5 Obtención de resultados	22
3.6 Despliegue	23
4. Reportes en tiempo real sobre status de proyectos	28
4.1 Entendimiento de requerimientos	30
4.2 Comprensión y obtención de los datos	31
4.3 Preprocesamiento de los datos	33
4.4 Análisis	35
4.5 Obtención de resultados	37
4.6 Despliegue	42
5. Resultados	44
5.1 Automatización de proceso de notificaciones	44
5.2 Reportes en tiempo real sobre status de proyectos	48
6. Conclusiones	51
Bibliografía	53

Resumen

El presente trabajo muestra el desarrollo de dos sistemas, el primero enfocado en la automatización de procesos de envío de correos electrónicos y el segundo se desarrolló principalmente en el análisis de datos para poder obtener visualizaciones generales y particulares del estatus y seguimiento de proyectos, información de arquitectura, infraestructura, entre otros.

Para el primer sistema la situación era que de forma manual se revisaba un archivo control, el cual contenía la información de los servicios, como su costo, duración, empresa responsable, solicitante del servicio, entre otros detalles, posteriormente, se generaban correos personalizados con información del nombre del servicio, el número de contrato, la fecha de finalización, entre otros y de esta manera informar al responsable para saber si va a querer o no renovar dicho servicio. El objetivo principal era automatizar el proceso y con ello, ahorrar el mayor tiempo posible. Se utilizaron las herramientas con las que contaba la empresa, en este caso *Google Workspace*¹. Con ayuda de *Google Sheets* se logró hacer una recolección, limpieza, procesamiento y análisis de los datos, para obtener una información procesada y utilizando *Google Script* se generaron una serie de códigos con los que se logró leer la información final y generar los correos electrónicos de manera automática y enviarlos. Con esto se cumplía el objetivo añadiendo la ventaja de que no fue necesario solicitar nuevos recursos adicionales e incluso instalaciones fuera de los sistemas de la empresa.

Para el segundo sistema, la empresa ya tenía información en formatos de *Google Sheets*, sin embargo, únicamente los visualizaban como forma de tabla, este era un problema pues la información no estaba actualizada al día y era difícil de ver grandes cantidades de las misma, el objetivo era crear tableros y dashboard los cuales ayudarán a visualizar mejor el comportamiento de la información. Mediante herramientas como *Google Sheets*, lenguajes de programación como Python, se realizaron las actividades de extracción, limpieza, procesamiento y análisis de datos, al final, se obtuvo un solo archivo y con él se utilizaron herramientas para visualización. Con esto se cumplió el objetivo de poder tener una mejor visualización de la información de la empresa, para así poder tomar decisiones en base tanto a datos históricos como datos actualizados en tiempo real.

¹ Sitio Google Workspace (<https://workspace.google.com/>)

1. Introducción

El presente trabajo se realizó en el contexto de mi experiencia profesional como "Web Analysis & BA services" dirigido al desarrollo de sistemas enfocados al tratamiento de diferentes datos e información con los cuales se pueda beneficiar la empresa donde trabajo. El análisis de datos es un proceso que conlleva el entendimiento de las necesidades de la empresa, obtención y preparación de datos, modelado y, por último, implementación, con lo cual se logra resaltar información útil e importante para la toma de decisiones y tener una mejor visualización del panorama de la empresa. Durante mi trayectoria en la empresa desarrollé distintos proyectos sobre datos, implementando un desarrollo por fases para la automatización de procedimientos, además obtener información del panorama general del desarrollo, testeo y monitorización de las aplicaciones en la empresa y así tener un mejor entendimiento de la situación para la toma de decisiones.

En el presente informe profesional muestro las actividades principales que desarrollé en la empresa, las cuales fueron el entendimiento de las necesidades de la empresa, la recopilación y preparación de los datos para así automatizar procesos en relación con requisiciones y envío de correos, además de la creación de informes o *dashboards* para la visualización de cómo se encuentra el desarrollo, testeo y monitorización de las distintas aplicaciones, todo este desarrollo se llevó a cabo en distintas fases y con la ayuda de las herramientas proporcionadas por la plataforma en la nube de *Google Workspace*².

A continuación, en las siguientes secciones se presentan los objetivos y la estructura del informe profesional.

² Sitio Google Workspace (<https://workspace.google.com/>)

1.1 Objetivos

Uno de los objetivos del análisis de datos es rescatar información importante o de utilidad para tener una mejor perspectiva de un tema en específico y así beneficiar el criterio a la hora de toma de decisiones. La empresa tenía dos necesidades principales, la primera era la automatización de un proceso, el cual requería la revisión de un gran volumen de información de requisiciones, para después con relación a 4 criterios establecidos se debía decidir cuándo mandar un correo con información específica. La otra necesidad era el obtener informes mediante *dashboard* (tablero interactivo) los cuales ayudarían a tener una mejor visualización de cómo se encontraba el desarrollo, testeo y el monitoreo de visitas a las diferentes aplicaciones que se tienen. Para solucionar ambos casos desarrollé análisis de datos, scripts y creé *dashboards* implementando las tecnologías de la nube de Google, como son, Google Sheets, Google Scripts, Google Colab y Data Studio de Google todo garantizando que fueran sistemas que se mantuvieran actualizados en tiempo real.

Hoy en día el desarrollo y avance en la tecnología va a un ritmo impresionante, esto ha desarrollado en gran parte la recolección de información y al mismo tiempo se ha impulsado el uso de análisis de toda la información recolectada. Las empresas han visto las ventajas y beneficios que se tiene al realizar estas actividades. La empresa donde laboré para la realización de este trabajo ha comenzado a implementar análisis de datos y automatización de procesos; su principal objetivo es tener un mejor control y toma de decisiones de lo que ocurre con los desarrollos, compras y un mejor control de versiones en relación con todo el tema de arquitectura.

1.2 Estructura del informe

En el capítulo dos hablo sobre lo que es el análisis de datos y las herramientas, además de explicar los proyectos en los que estuve trabajando. Para el capítulo tres, explico el desarrollo del proyecto de automatización. En el capítulo cuatro presento el segundo desarrollo en relación con la generación de *dashboards*. En el capítulo cinco se muestran los resultados obtenidos, así como sus análisis. Y por último, en el sexto capítulo se dieron las conclusiones sobre todo este trabajo.

2. Análisis de Datos y herramientas

Debido a los avances tecnológicos que se han desarrollado en los últimos años en relación con el almacenamiento y recolección de información, el análisis de los mismos se ha vuelto una actividad fundamental para las empresas, pues con los resultados que se tienen se utilizan para tener un mejor panorama de una situación en específica y poder tomar mejores decisiones o acciones que ayuden a lograr los objetivos planteados por la empresa. Además, la industria ha comenzado a automatizar sus procesos y por lo mismo necesita modelos de calidad, para ello se apoya en la minería y analítica de datos, y con ello optimizar de manera eficiente tareas o procesos que requieran mucho tiempo de manera manual.

El análisis de datos se llega a definir como un proceso por el cual los datos pasan por una serie de acciones de limpieza, transformación y modelado para así descubrir información útil para la toma de decisiones (Hegde, 2022). El propósito fundamental del análisis de datos es extraer la mayor cantidad de información útil y con ello tomar una decisión basada en el análisis y conclusiones de los mismos.

Actualmente, debido a la gran cantidad de información que se encuentra disponible, los datos provienen de diferentes fuentes y estructuras, por lo cual se definen en diferentes tipos (Maryville University, 2021):

- Big data. Se define como un gran volumen de datos el cual continúa creciendo a un ritmo exponencial a lo largo del tiempo.
- Datos estructurados / no estructurados.
 - Datos estructurados son aquellos que tienen un modelo de datos definido, como una base de datos, tablas, etc.
 - Los datos no estructurados no están definidos en filas y columnas, además pueden incluir diferentes tipos de formatos como videos, fotos, audio, texto, entre otros.
- Metadatos. Son una forma de datos los cuales describen y dan información sobre otros datos. Esto nos permite organizar datos no estructurados en categorías o clasificaciones.

- Datos en tiempo real. Este tipo de datos son de gran utilidad cuando las decisiones requieren de tener la información más actualizada. Para este tipo de casos el uso de la nube es muy común.
- Datos de la máquina. Gracias a la Internet de las cosas (IoT), los sensores y otras tecnologías, los sistemas de fábrica y otras máquinas, la tecnología de la información y la infraestructura de telecomunicaciones, los automóviles inteligentes, los dispositivos portátiles y más pueden generar datos automáticamente. Este tipo de datos se conoce como datos de máquina porque son producidos en su totalidad por máquinas sin instrucción humana.

Los procedimientos y técnicas implementadas en el análisis de datos son de ayuda para obtener información relevante en los datos, como métricas, hechos y cifras. La primera técnica de análisis se realiza mediante datos cualitativos y la segunda con datos cuantitativos. Estas técnicas se pueden utilizar de forma separada o en conjunto con otras, las cuales ayudan a las empresas y los responsables de la toma de decisiones ya que adquirieren conocimientos de las situaciones a partir de la información histórica que tienen sin importar el tipo de datos (Maryville University, 2021).

2.1 CRISP-DM

Para el análisis de datos que empleé en la empresa, utilicé la metodología de CRISP-DM, la cual se divide en seis fases (Inc & SPSS Inc, 2000):

I. Entendimiento de los requerimientos del cliente.

Aquí el enfoque es comprender los objetivos del proyecto y así poder diseñar un plan preliminar para alcanzar dichos objetivos.

II. Comprensión de los datos.

En esta fase se realiza la recolección de datos y el estudio preliminar de los mismos, lo cual permite identificar problemas, qué tipo de datos y con qué calidad se tienen para su análisis posterior.

III. Preprocesamiento de los datos.

En este punto se ejecutan actividades de limpieza para tener un conjunto o dataset óptimo para su procesamiento, buscando tener homogeneidad, identificar datos nulos, incompletos, integración de las diferentes fuentes y poder organizarlos en una estructura definida (tablas, base de datos, entre otros).

IV. Análisis de datos

Aquí se seleccionan y aplica alguna técnica que se adapte mejor al procesamiento de los datos que se tienen y poder obtener información relevante para la empresa. Cada técnica requiere una estructura específica de los datos. Por lo mismo, puede regresar a la fase anterior.

V. Obtención de resultados.

En esta fase se evalúan y revisan los resultados obtenidos con el objetivo del proyecto. Para posteriormente tomar una decisión sobre cuál o cómo será la presentación de los resultados del proceso de análisis de datos.

VI. Despliegue.

Esta fase puede ser simple como la creación de *dashboard* o incluso más compleja como la integración de programas para su uso final, dependerá del requerimiento del cliente.



Figura 2.1 Metodología de análisis de datos

2.2 Análisis de datos en las Empresas

Hoy en día debido a los diferentes cambios que han sufrido las empresas, como por ejemplo, la migración a la nube, el *home office*, etc, las tomas de decisiones se han vuelto mucho más complejas y de mayor importancia, por lo mismo las actividades de análisis de datos, minería de datos e incluso ciencia de datos, han sufrido un incremento en las responsabilidades que tienen en las empresas.

De acuerdo a IBM, el 90% de los datos que se tienen hoy en día se han creado y almacenado en los últimos 10 años, (Redacción, 2020), esto quiere decir, que las empresas comienzan a digitalizarse a un ritmo más acelerado y el manejo de toda esta información se ha vuelto una actividad crítica y fundamental para las empresas. La mayoría de las empresas no estaba preparada para este tipo de actividades, aunque almacenaban grandes cantidades de información no contaban con las herramientas necesarias para el procesamiento y análisis de la misma.

Actualmente existen muchas herramientas y software para estas actividades, sin embargo, no todas son gratuitas, muchas de ellas requieren del pago de licencias, una opción para solucionar este problema es, en primer lugar, conocer las diferentes herramientas con las que cuenta la empresa y capacitar a los analistas para hacer uso de las mismas, esto es de suma importancia pues el correcto aprovechamiento y explotación de los datos pueden tener grandes impactos en el crecimiento de una empresa, llegando incluso a incrementar en 60% los márgenes de operación. (Jacob Escobar, 2021).

Debido a que las empresas siguen aprendiendo a utilizar sus datos de manera correcta, es importante, que no se cierren al mundo exterior, pues esto los ayudará a tener una mejor perspectiva y conocimientos actuales del mundo del análisis de datos. Por último, para una empresa no solamente es importante realizar un buen análisis sino también el cómo mostrarlo, ya que con frecuencia, los resultados son presentados a juntas directivas y para este fin es indispensable tener una buena presentación de las conclusiones obtenidas luego de haber realizado todo el proceso de análisis de datos. (Souza, 2021).

2.3 Data Studio

Data Studio es una herramienta gratuita que convierte sus datos en informes y paneles claros, totalmente personalizables y fáciles de consultar y compartir. Ayuda a transformar los datos en reportes, informes o *dashboards* interactivos y claros. Estos son completamente personalizables en estructura, color, diseño, entre otros. Además, permite compartir, en la forma que se desea, los desarrollos con la ventaja de que no sea necesaria ningún tipo de instalación, únicamente se requiere de tener alguna cuenta de Google activa, en la Figura 2.2 se presenta la interfaz principal de Data Studio. (Google, 2022-b)

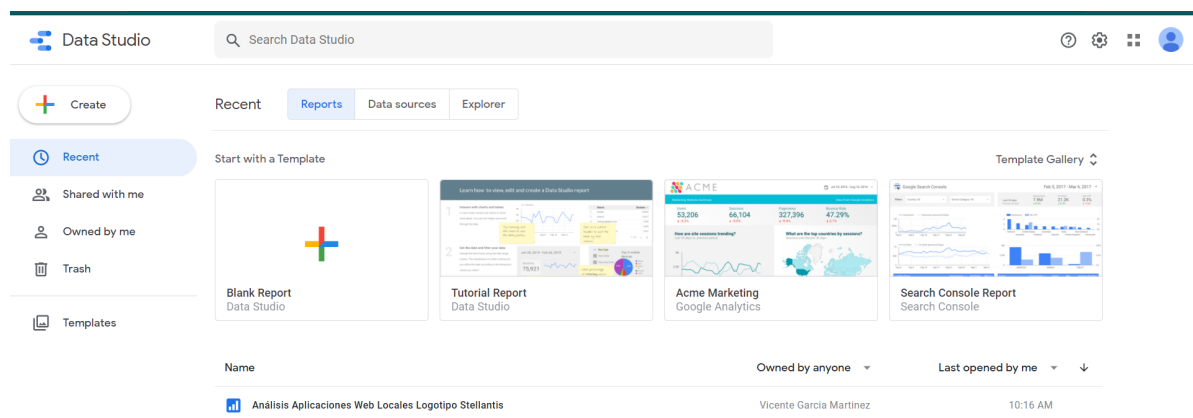


Figura 2.2 Interfaz de *Google Data Studio*.

Entre sus características más importantes es que a través de los datos que le conectemos se pueden mostrar mediante diferentes gráficos como de barras, líneas, mapas geográficos, circulares, áreas, tablas dinámicas, entre otros, en adición, una de las funciones más útiles para generar *dashboards* es el uso de filtros o controles de periodo, que mediante ellos se puede controlar los datos que son mostrados, logrando personalizar o especificar los reportes en tiempo real, haciendo que sean específicos para algún caso o enfoque definido.

Con *Data Studio* se puede conectar una gran variedad de fuentes, esto garantiza que los reportes serán actualizados en tiempo real siempre y cuando los archivos se encuentren en línea, dentro de sus conexiones más utilizadas están *Google Analytics*, consumir Hojas de cálculo como Excel o Google Sheets e incluso la integración o conexión a bases de datos como *BigQuery*, *MySQL* o *PostgreSQL*.

2.4 Google Apps Script

Google Apps Script es una plataforma de desarrollo de aplicaciones rápida que agiliza y facilita la creación de aplicaciones comerciales que se integran con *Google Workspace* (Google, 2022-a), es decir, se conecta con los servicios de Google como por ejemplo Gmail, Google Sheets, Docs, entre otros. Las aplicaciones se escriben con código en JavaScript moderno, este ya incluye las librerías necesarias para las conexiones con *Google Workspace*, es decir, ya no existe la necesidad de instalar nada, tanto para su desarrollo como su ejecución, adicionalmente desde el sitio oficial se proporciona un editor de código; la Figura 2.3 muestra el IDE que proporciona Google. Por último, los scripts o aplicaciones que se desarrollan se ejecutan en los servidores de Google.

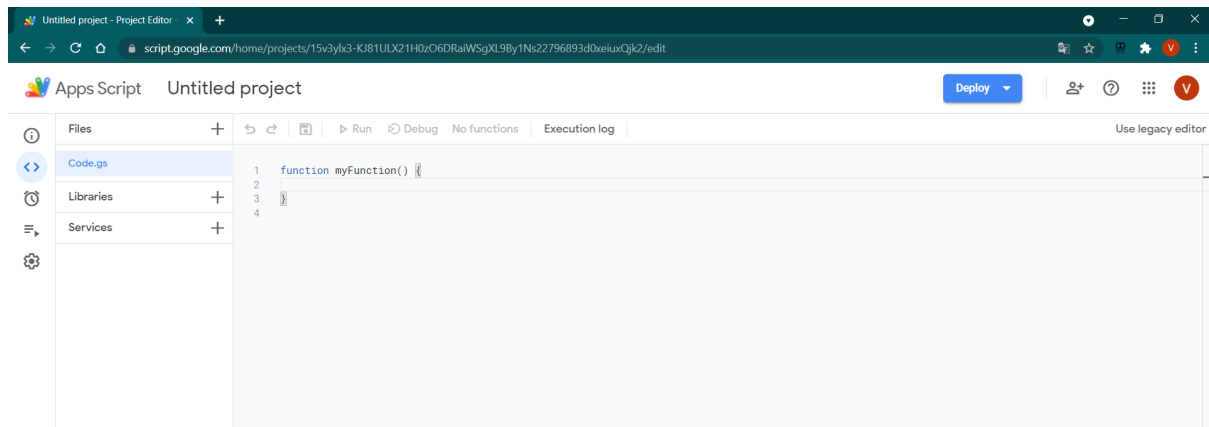


Figura 2.3 IDE de *Google Apps Script*.

Dentro de las funciones que se pueden desarrollar en *Google Apps Script* están las siguientes:

- Agregar botones dentro de los menús y barras laterales en las aplicaciones de Google Docs, Sheets y Forms.
- Escriba funciones y macros personalizadas las cuales se ejecutarán en Google Sheets.
- Interactúe con los servicios de Google , como Analytics, Calendar, Drive, Gmail y Maps.
- Optimizar flujos y procesos de trabajo de Google Chat mediante la creación de un chat bot.

2.5 Desarrollos en la empresa

Un objetivo que tenía la empresa desde 2020, fue optimizar procesos que involucraron mucho tiempo debido a que en su mayoría todo se hacía de manera manual y, por otro lado, conocer el estado de diferentes proyectos para una mejor toma de decisiones. Estos desarrollos se dividieron en dos secciones.

El primer proyecto se enfocó en el análisis de una tabla de información con requisiciones, cuyo objetivo fue saber en qué momento notificar mediante correo electrónico de manera automática al usuario o empresa que su servicio estaba próximo a vencer. El segundo proyecto se basó en la recolección de información de distintos proyectos y su procesamiento para presentar *dashboard* con información relevante para la empresa y, con ello, poder tomar decisiones en beneficio del mismo, además de llevar un mejor control del status de las diferentes actividades que se tienen.

Las soluciones que se necesitaron se desarrollaron con herramientas de programación como el uso de *Google Colab* para el procesamiento de los datos, *Google Script* para la programación del envío automático de correos electrónicos mediante Gmail para la notificación sobre las requisiciones y además de utilizar interfaces para el desarrollo de *dashboard* como son *Data Studio* de Google y *Qlik Sense*. Las herramientas e interfaz de Google fueron de gran importancia ya que permiten que el sistema pudiera ser utilizado en tiempo real, esto debido a que las fuentes principales de información se encuentran en constantes cambios y/o actualización y era de suma importancia para el negocio que esos cambios se vieran automáticamente reflejados.

3. Automatización de proceso de notificaciones

La empresa cuenta con servicios, los cuales están en constante renovación, por lo cual era de suma importancia llevar el control sobre lo relacionado con la requisición del servicio. Para esto el cliente tenía un archivo en la nube, específicamente en Google Drive como se muestra en la Figura 3.1, con toda la información necesaria para poder identificar la requisición del servicio, así como características importantes como son la fecha inicio y fin del contrato, costo total, contacto del proveedor y empresa.

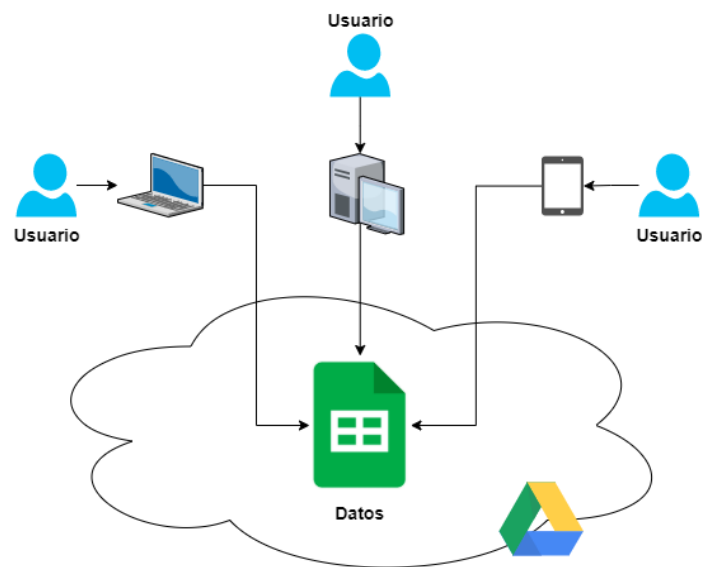


Figura 3.1 Interacción del usuario con los datos.

Con esta información, el cliente le dedicaba de 3 a 5 días a la revisión de la misma para posteriormente seleccionar y filtrar los campos necesarios y así poderlos compartir con otra integrante de su equipo. Esta persona era la encargada de estructurar los correos que eran mandados a cada uno de los clientes. Esto conllevaba ciertos problemas, pues la revisión que se hacía de los datos se efectuaba de manera manual y era propensa a fallos debido a la cantidad de datos que se contenían en el archivo. Por otro lado, el envío de los correos también era de forma manual, es decir, se creaba uno por uno con la información proporcionada, por lo cual también requería de bastante tiempo.

El objetivo del proyecto fue optimizar el proceso mencionado de tal manera que se lograra realizar de forma automática, aprovechando las herramientas en la nube y sin tener que realizar cambios en la forma en cómo se estaba recolectando la información. Para esto, se comenzó con el desarrollo de la metodología de datos en sus diferentes fases, las cuales se explicarán a continuación.

3.1 Entendimiento de requerimientos

Como primer paso para el desarrollo de este proyecto, se llevaron a cabo reuniones por medio de Google Meet, en las cuales se me proporcionó el detalle del estado actual y los objetivos a los que querían llegar.

El sistema que se solicitó debía cumplir ciertas características, la primera es que pudiera obtener de manera automática qué requisiciones se tenían que renovar siguiendo los siguientes criterios:

1. Si el servicio cuesta menos de 20,000 dólares, se debe notificar de la renovación del servicio antes de 2 meses que finalice el contrato.
2. Si el servicio cuesta entre 20,000 y 100,000 dólares, se debe notificar de la renovación del servicio antes de 3 meses que finalice el contrato.
3. Si el servicio cuesta más de 100,000 dólares, se debe notificar de la renovación del servicio antes de 4 meses que finalice el contrato.

Adicionalmente, la notificación se debía realizar mediante correo electrónico con base en los datos que se tengan en el registro de la requisición, es decir, dicho correo debía contener la información de qué requisición estaba a punto de vencer, así como un folio, el cual fue el identificador de la misma, las fecha final del servicio y a qué empresa correspondía. Este correo se tenía que mandar al responsable del servicio y con copia a los diferentes grupos de seguimiento. Todo este proceso se solicitó que se ejecutará una vez al mes.

Con base a lo anterior, se realizó un análisis para decidir si con las herramientas actuales eran suficientes. Esto arrojó que no habría problema en desarrollar la solución y a partir de esto se procedió a la siguiente fase.

3.2 Comprensión y obtención de los datos

Para esta segunda fase se me proporcionó acceso al archivo llamado “GPV 2020 CONTROL DE G&A”, el cual se presenta una captura en la Figura 3.2. En este archivo contiene toda la información sobre las requisiciones que se encuentran activas desde el año 2020 a la fecha. Se me explicó cada uno de los campos para así poder entender qué información sería de utilidad para el desarrollo del proyecto.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	# RQ	PO	# RQ	RQ PESOS	RQ USD	IC	Mo	STATUS	Servicio Control	Servicio	DESCRIPCION	DESCRIPCION CORTA			
349	347	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Jun
350	348	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Rol
351	349	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Rol
352	350	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Vict
353	351	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Ton
354	352	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Jun
355	353	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Rol
356	354	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Rol
357	355	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Vict
358	356	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Ton
359	357	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Jun
360	358	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Arr
361	359	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Edu
362	360	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Mig
363	361	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Mig
364	362	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Arr
365	363	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Arr
366	364	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Arr
367	365	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Arr
368	366	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Will
369	367	R				21.60	MON		Servicios adquiridos	Servicios adquiridos					Mig
370	368	R				19.00	USD		Servicios adquiridos	Servicios adquiridos					Arr
371	369	R				19.00	USD		Servicios adquiridos	Servicios adquiridos					Rol
372	370	R				19.00	USD		Mantenimiento	Mantenimiento					Rol

Figura 3.2 Archivo “GPV 2020 CONTROL DE G&A”.

Para el análisis de los datos y poder obtener la información necesaria se llegó a la conclusión de que las columnas necesarias y relevantes fueron las siguientes:

- # RQ : Corresponde al número de la requisición, este valor es único por lo que se puede tomar como ID.
- PO : En caso de tener un registro anterior corresponde al número de la requisición previa.

- **PROVEEDOR** : Nombre del proveedor al que se le compra o contrata el servicio o producto.
- **DESCRIPCIÓN** : Breve explicación del entendimiento y características del producto o servicio.
- **RESPONSABLE** : Nombre de la persona encargada de la gestión y control de la requisición.
- **INICIO** : Fecha de inicio de la contratación.
- **TERMINO** : Fecha de término de contrato con el servicio contratado.
- **O.C. PESOS** : Precio de la requisición en pesos mexicanos.
- **O.C. USD** : Precio de la requisición en dólares americanos.
- **STATUS** : Este valor nos indicaba si la requisición fue aceptada o cancelada.

3.3 Preprocesamiento de los datos

Debido a que únicamente se necesitó el análisis de las requisiciones que estaban previamente registradas en SAP se hizo la limpieza dejando únicamente los datos o registros que tuvieran “# RQ”, un ejemplo se muestra en la Figura 3.3, con lo cual se garantiza el registro previo y además cumplieran con la condición de que hayan sido aceptadas en el proceso; esto se verificó utilizando el valor de STATUS mediante el cual se eliminaron los registros que no tuvieran el valor de “Colocada”.

#	Itrac	R/D	PO	# RQ	RQ PESOS	RQ USD	TC Base	Moneda	STATUS
284		R					0	19.00	USD
285		R					4	19.00	USD
286		R					0	19.00	USD
287		R					0	19.00	USD
288		R					0	19.00	MXN
289		R					4	19.00	MXN
290		R					1	19.00	USD
291		R					1	19.00	USD
292		R					6	19.00	USD
293		R					0	19.00	USD
294		R					4	19.00	MXN
295		R					0	19.00	USD
296		R						19.00	USD
297		R						19.00	USD
298		R						19.00	USD

Figura 3.3 Archivo con los valores a filtrar.

Estas dos condiciones garantizó que los demás campos los tuvieran con valores y así lograr tanto el análisis exploratorio como la generación del script. Este tipo de limpieza se realizó mediante funciones tipo *FILTER* en el software de *Google Sheets*, la decisión de optar por usar esta interfaz la tomé para poder conservar la misma estructura que ellos implementan, además, de poder garantizar que siempre estaría actualizada, pues al ser una herramienta en línea, permite obtener los cambios en tiempo real.

3.4 Análisis

Uno de los principales retos para el desarrollo del análisis fue la selección de la herramienta con la cual se desarrollaría, la principal barrera se encontraba debido a las herramientas con las que la empresa contaba, en ese tiempo, la empresa carecía de herramientas enfocadas para análisis, por lo mismo se optó por utilizar *Google Sheets*, esto permitía poder seguir usando sus fuentes de datos de manera casi transparente, para esto se utilizaron diferentes funciones.

1. *importrange*, que me permitió extraer la información específica de sus diferentes archivos control.
2. *if*, mediante esta función se implementó para poder obtener, mediante cálculos y comparaciones, cuál era el caso al cual correspondía cada una de las requisiciones.

El primer dato que se necesitaba obtener fue el cálculo de cuántos meses faltan para el vencimiento de las diferentes requisiciones, al ser una gran cantidad de requisiciones a analizar se implementaron funciones *if* anidadas, figuras 3.4 y 3.5.

```
=IF(ISERROR((IF(B5 = "", "SIN TERMINO", IF((DATEDIF(TODAY(), B5, "M")) > 1, (DATEDIF(TODAY(), B5, "M")), "Menos de 2 meses")))), "VENCIDA", IF(B5 = "", "SIN TERMINO", IF((DATEDIF(TODAY(), B5, "M")) > 1, (DATEDIF(TODAY(), B5, "M")), "Menos de 2 meses")))
```

Figura 3.4 Función *if* para obtener el número de meses que faltan para vencimiento.

```
=IF(D3="SIN TERMINO", "SIN TERMINO", IF(D3="VENCIDA", "VENCIDA", IF(C3 <= 20000 , IF(D3 = 2 , "Caso 1", "Otro"), IF(C3 <= 100000 , IF(D3 = 3 , "Caso 2", "Otro"), IF(C3 > 100000 , IF(D3 = 4 , "Caso 3", "Otro"), "Otro")))))
```

Figura 3.5 Función *if* para determinar en el caso al cual aplican.

En esta función no solo se utilizaron condicionales, sino que dentro de la misma se realizaron diferentes cálculos para obtener el valor específico de los meses faltantes, para poder explicar las funciones anidadas se presenta, a continuación, la definición específica de las funciones generales.

1. *ISERROR()* : esta función nos permitió atrapar algún error, o datos fuera de rango, a la hora de leer el dato de la fecha de vencimiento y determinar si ya se había vencido o no.
2. *DATEDIF()* : mediante esta función, se nos permitió obtener el número de meses que existe entre dos fechas específicas.
3. *TODAY()*: función para obtener la fecha del día actual en el que se utilice la hoja de *Google Sheets*.

Una vez explicadas las funciones se muestra el detalle de función utilizada:

IF((DATEDIF(TODAY(),B50,"M"))>1,(DATEDIF(TODAY(),B50,"M"),"Menos de 2 meses")))), mediante esta función, se comprobó el si una requisición entraba en una excepción, es decir, si su renovación del servicio entraba en un rango menor a dos meses, en caso de ser mayor se colocaba el número de meses faltantes.

IF(D3="SIN TERMINO", "SIN TERMINO", IF(D3="VENCIDA", "VENCIDA", IF(C3 <= 20000 , IF(D3 = 2 , "Caso 1", "Otro"), IF(C3 <= 100000 , IF(D3 = 3 , "Caso 2", "Otro"), IF(C3 > 100000 , IF(D3 = 4 , "Caso 3", "Otro"), "Otro"))))), esta función se utilizó para determinar en cuál de los casos, definidos en la sección de entendimiento de requerimientos, entraba cada una de las requisiciones.

IFS(L18="Caso 1", "TRUE", L18="Caso 2", "TRUE", L18="Caso 3", "TRUE", L18="Otro", "FALSE", L18="SIN TERMINO", "FALSE", L18="VENCIDA", "FALSE"),

función implementada para determinar, mediante un valor *True* o *False*, si la requisición entraría en el proceso de notificaciones automatizadas.

Mediante el uso de las diferentes fórmulas se logró determinar que requisiciones se necesitaban notificar y cuales no, figura 3.6, y con ello, únicamente pasar la información necesaria al script de envío de notificaciones.

TERMINO	RQ USD	Calculo meses a vences	Casos	T/F
09/Sep/2022	17,621.05	2	Caso 1	TRUE
30/Sep/2022	28,118.13	3	Caso 2	TRUE
06/Nov/2022	141,926.40	4	Caso 3	TRUE
31/Dec/2019	56,690.53	VENCIDA	VENCIDA	FALSE
	15,789.47	SIN TERMINO	SIN TERMINO	FALSE
30/Jun/2022	1,551,294.73	Menos de 2 meses	Otro	FALSE

Figura 3.6 Resultados de aplicar las diferentes funciones.

Para tener una mejor visualización del comportamiento y estatus de las requisiciones a lo largo del tiempo se desarrolló un *dashboard*, el cual se presenta en la Figura 3.7, con el cual se logró interactuar en tiempo real y así poder visualizar aspectos de importancia para la empresa.

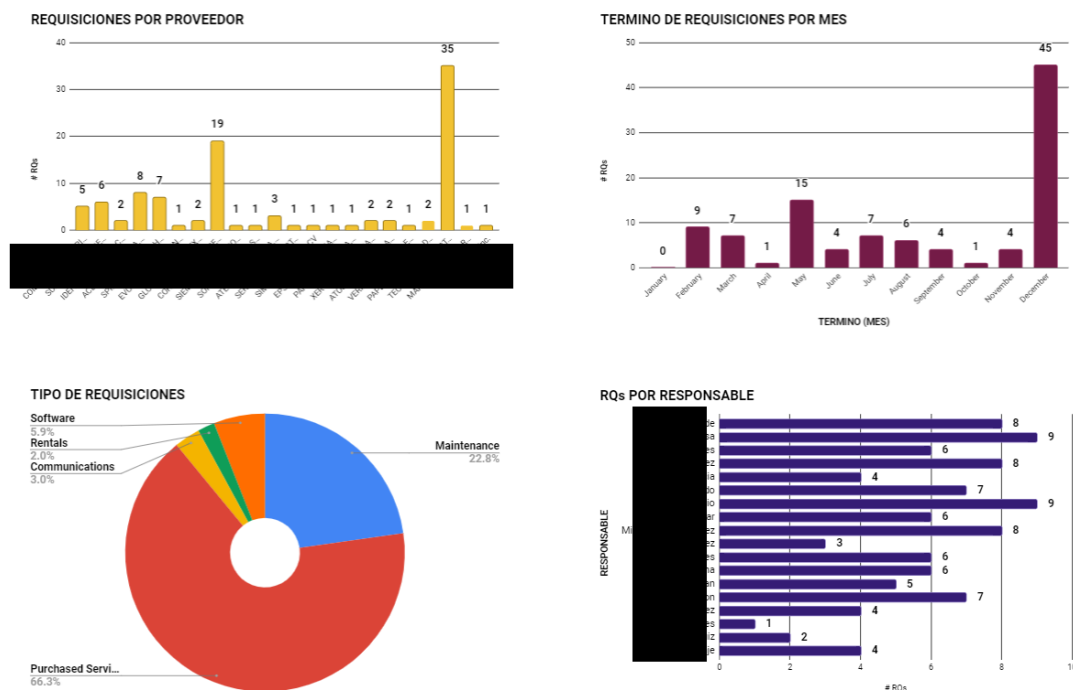


Figura 3.7 *Dashboard* de Requisiciones General. Se ocultaron datos confidenciales.

El objetivo de este análisis fue ver el comportamiento de las requisiciones sobre cuatro aspectos relevantes, el primero es el proveedor, con esta información se logró visualizar cuál empresa estaba siendo requerida más, el segundo son el término de requisiciones por mes, con lo cual se pudo tomar una decisión estratégica para considerar el flujo de trabajo que se tendrá en este caso a finales de año. El tercer aspecto fue el tipo de requisiciones, este dio un panorama general del tipo de servicios que se están contratando y así decidir futuros presupuestos específicos, por último se realizó una visualización de cuantas requisiciones tenían asignados los diferentes responsables y así saber la situación de cómo se estaba distribuyendo las responsabilidades.

Siguiendo con la idea de conservar la interfaz de *Google Sheets*, opté por utilizar tablas pivote con las cuales a través de paneles de configuración fue posible crear gráficas dinámicas. Esto ayuda a poder tener y seleccionar información específica para saber la situación actual de los datos y su comportamiento.

Para poder interactuar con el *dashboard* se creó una sección mediante la cual se podía seleccionar el año, proveedor y tipo de requisición, dando como resultado el número total de requisiciones que cumplen los criterios seleccionados y mostrando a la vez las gráficas explicadas anteriormente; con esto se logró visualizar casos en específico para una mejor toma de decisiones. Este se muestra en la Figura 3.8.

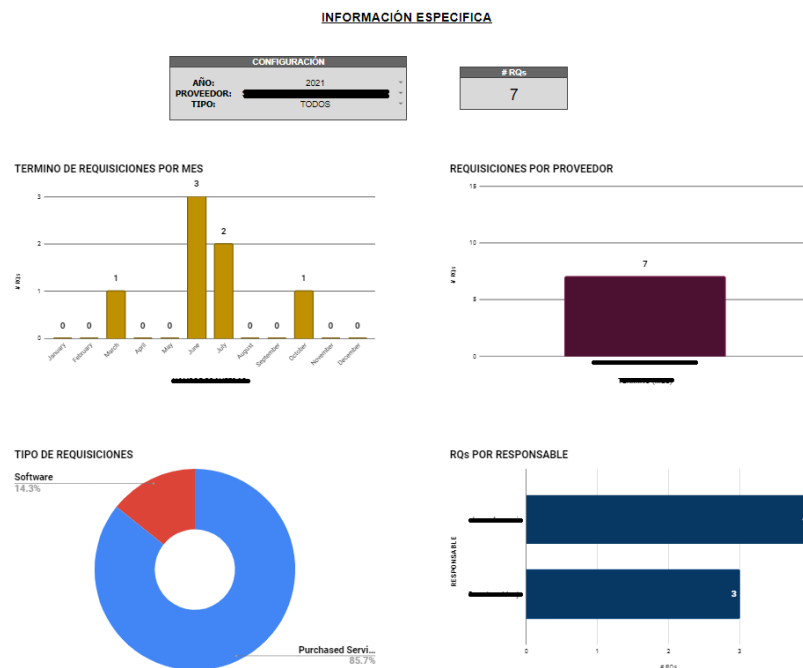


Figura 3.8 *Dashboard* de Requisiciones Específicas. Se ocultaron datos confidenciales.

Para poder la información a la mano ya filtrada, se agregó una tabla dinámica, Figura 3.9, donde únicamente se mostraba el detalle de las requisiciones que cumplen los criterios.

#	fin	litrac	R/D	PO	# RQ	RQ PESOS	RQ USD	TC Base	Moneda	STATUS	DESCRIPCION	DESCRIPCION CORTA	RESPONSABLE
22	0 t		R										
43	3 t		R										
44	1 t		R										
45	3 t		R										
56	4 t		R										
123	4 t		R										
153	4 t		R										

Figura 3.9 Detalle de Requisiciones Específicas. Se ocultaron datos confidenciales.

3.5 Obtención de resultados

Para esta fase se realizó una revisión del análisis previo, con lo cual se verificó que se estuvieran tomando correctamente las requisiciones considerando los filtros aplicados en la limpieza de datos y ser guardados en un archivo nuevo llamado “Servicio próximo a vencer”, se muestra en la Figura 3.10, a partir de este se ejecutarán los programas siguientes. Posteriormente, se tomó la decisión de continuar con el script para la automatización del proceso de notificaciones.

# RQ	PO	PROVEEDOR	DESCRIPCION	Approval SC	RESPONSABLE INICIO	TERMINO	O.C. PESOS	O.C. USD
						01/Aug/2021	31/Jul/2021	
						01/Jul/2021	30/Jun/2021	
						01/Jul/2021	30/Jun/2021	
						01/Jun/2021	31/May/2021	
						01/Jun/2021	31/May/2021	
			Servicio de mantenimiento preventivo/correctivo a equipos			01/Jun/2021	31/May/2021	
						01/Jun/2021	31/May/2021	
						01/Jun/2021	31/May/2021	

Figura 3.10 Detalle para revisión de requisiciones. Se ocultaron datos confidenciales.

Para tener una mejor visualización se desarrolló una extracción de datos con las condiciones correspondientes y dar un formato en tabla, se muestra en la Figura 3.11, este posteriormente sirvió para que fueran los datos de entrada para el script de notificación automática.

Name	Message1	Parametro1	Message2	Parametro2	Message3	Parametro3	Vencimiento
J	DESCRIPCIÓN;		PROVEEDOR;		S.A. OC ACTUAL;		
	DESCRIPCIÓN;		PROVEEDOR;		DE OC ACTUAL;		
2	DESCRIPCIÓN;		PROVEEDOR;		OC ACTUAL;		
	DESCRIPCIÓN;		PROVEEDOR;		OC ACTUAL;		
e	DESCRIPCIÓN;		PROVEEDOR;		NG OC ACTUAL;		
9	DESCRIPCIÓN;		PROVEEDOR;		A OC ACTUAL;		
z	DESCRIPCIÓN;		PROVEEDOR;		OC ACTUAL;		
5	DESCRIPCIÓN;		PROVEEDOR;		OC ACTUAL;		

Figura 3.11 Formato de Requisiciones para Script. Se ocultaron datos confidenciales.

3.6 Despliegue

En esta fase se crearon dos script mediante el lenguaje de GS (Google Script) esto se decidió con base en los requerimientos de la empresa, ya que se contaba con todos sus sistemas de comunicación como chat, correos, reuniones, entre otros, y almacenamiento de archivos mediante la tecnología de Google Drive. El flujo de funcionamiento se muestra en la Figura 3.12.

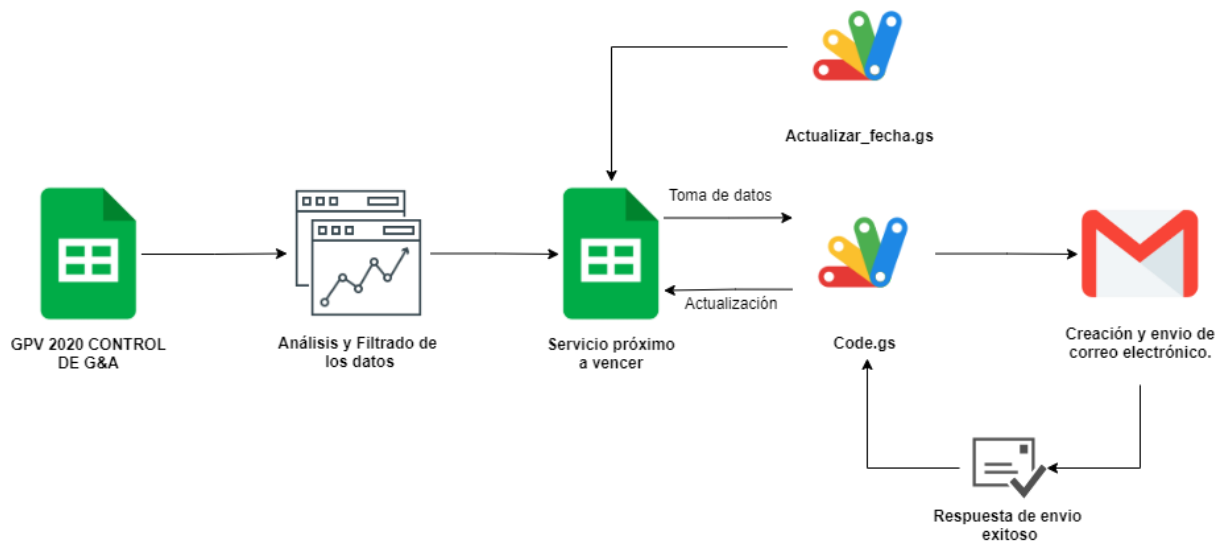


Figura 3.12 Flujo de trabajo del envío de notificaciones automatizado.

El primer script llamado “Actualizar_fecha.gs”, se muestra en la Figura 3.13, se empleó para la actualización de fechas de notificación en el sheets “Servicio próximo a vencer”. Este se utilizó como control de las notificaciones de las requisiciones. La principal función del script fue garantizar que sin importar el día en que fuera ejecutado el script principal, se mandarían correctamente los correos, mediante una verificación de cuántos meses faltan para que venciera la requisición.

```

Code.gs
Actualizar_fecha.gs.gs
Libraries
Services

1  /**
2  * Sends non-duplicate emails with data from the current spreadsheet.
3  */
4  function dates_update() {
5      var sheet = SpreadsheetApp.getActiveSheet();
6      var startRow = 2; // First row of data to process
7      var numRows = CountCol(); // Number of rows to process
8      var dataRange = sheet.getRange(startRow, 1, numRows, 13);
9      // Fetch values for each row in the Range.
10     var data = dataRange.getValues();
11     for (var i = 0; i < data.length; ++i) {
12         var date1 = new Date();
13         sheet.getRange(startRow + i, 12).setValue(date1);
14         // Make sure the cell is updated right away in case the script is interrupted
15         SpreadsheetApp.flush();
16     }
17 }
18
19 function onOpen(){
20     var spreadsheet = SpreadsheetApp.getActive();
21     var menuItems = [
22     {name: 'Dates Update', functionName: 'dates_update'}
23     ];
24     spreadsheet.addMenu('Dates Update', menuItems);
25 }
26
27 function CountCol(){
28     var sheet = SpreadsheetApp.getActiveSheet();
29     var data = sheet.getDataRange().getValues();
30     for(var i = data.length-1 ; i >=0 ; i--){
31         if (data[i][0] != null && data[i][0] != ''){
32             return i ;
33         }
34     }
35 }

```

Figura 3.13 Script para actualización de fechas.

El segundo script fue llamado “Code.gs”, se muestra en las Figuras 3.14 y 3.15, este realizaba la función principal el cual era obtener los datos correspondientes como fueron destinatarios, nombre del proveedor, descripción del servicio y número de orden OC. Esta consulta de datos se efectuaba mediante la conexión entre el script y el sheets de “Servicio próximo a vencer”.

Code.gs	3	*/
Date.gs	4	function sendEmails() {
	5	var EMAIL_SENT = "EMAIL_SENT";
	6	var sheet = SpreadsheetApp.getActiveSheet();
	7	var startRow = 2; // First row of data to process
Libraries	8	var numRows = CountCol(); // Number of rows to process
	9	// Fetch the range of cells A2:G4
Services	10	var dataRange = sheet.getRange(startRow, 1, numRows, 13);
	11	// Fetch values for each row in the Range.
	12	var data = dataRange.getValues();
	13	for (var i = 0; i < data.length; ++i) {
	14	var row = data[i];
	15	var emailAddress = row[0]; // First column
	16	var name = row[1];
	17	var message1 = row[2]; // DESCRPCIÓN;
	18	var description = row[3]; // description of service
	19	var message2 = row[4]; // PROVEEDOR;
	20	var service = row[5]; // Name of service
	21	var message3 = row[6]; // OC ALCTUAL;
	22	var number = row[7]; // Number of service
	23	var Vencimiento = sheet.getRange(startRow + i, 9).getValues().toString();
	24	var notification = sheet.getRange(startRow + i, 13).getValues().toString();
	25	var notificationNumber = parseInt(notification).toFixed();
	--	

Figura 3.14 Obtención de datos mediante script.

Posteriormente, se realizaba la creación de todo el cuerpo del correo con los datos obtenidos previamente, agregando tanto el destinatario como el asunto de manera automática, adicionalmente se enviaban copias de estos correos a los gerentes, ya que lo solicitaron así, para esto fue suficiente con agregar sus correos al script.

```

var message4 = 'Buen dia ' + name + '\n\nEl ICT Spending Committee te informa que el siguiente servicio a tu cargo, está próximo a vencer y necesitamos tu intervencion para iniciar el proceso de renovacion.\n\n' + message1 + '\n' + message2 + '\n' + message3 + '\n' + message4 + '\n' + Vencimiento;
var message = message4 + '\n\nFavor de continuar con el proceso de adquisiciones ingresando tu requerimiento en la siguiente liga: ' + Vencimiento;
//\n\nTambien te agradecemos nos notifiquemos (a los copiados en este correo), en caso de que este servicio ya NO deba renovarse';

var emailSent = sheet.getRange(startRow + 1, 11).getValues().toString();// Tenth column

if(notificationNumber >= 1){
  if (emailSent != "EMAIL_SENT") { // Prevents sending duplicates
    var subject1 = 'Importante! Servicio proximo a vencer';
    MailApp.sendEmail(
      {
        to: emailAddress,
        subject: subject1,
        body: message,
        cc: ' '
      }
    );
    sheet.getRange(startRow + 1, 11).setValue(EMAIL_SENT);
    // Make sure the cell is updated right away in case the script is interrupted
    SpreadsheetApp.flush();
  }
}
}

```

Figura 3.15 Script para envío de correos. Se ocultaron datos confidenciales.

Para lograr un control del envío de correos en el archivo de “Servicio próximo a vencer” se agregó una columna llamada “Email Send?”, se muestra en la Figura 3.16, esta se actualizaba a la hora que un correo fuera mandado de manera exitosa sobre la requisición correspondiente. Esto garantizó que sin importar que el script se corriera de forma manual o automática únicamente, se mandaran las notificaciones que aún no habían salido e incluso protegía de errores externos de conexión entre otros.

Message3	Parametro3	Vencimiento	Notification	Email Send ?	Current date	Days
OC ACTUAL;	61459464	31/Dec/2021	30/Sep/2021	EMAIL_SENT	8/31/2021	30
OC ACTUAL;	61459465	31/Dec/2021	30/Sep/2021	EMAIL_SENT	8/31/2021	30
OC ACTUAL;	61459466	31/Dec/2021	30/Sep/2021	EMAIL_SENT	8/31/2021	30
OC ACTUAL;	61459469	31/Dec/2021	30/Sep/2021	EMAIL_SENT	8/31/2021	30

Figura 3.16 Columna de verificación de envío de correos.

Para la ejecución de los scripts fue configurado el ambiente, el cual se ejecute el último día hábil de cada mes. Sin embargo, para casos extraordinarios o especiales se optó por agregar dos botones dentro del archivo “Servicio próximo a vencer”. Se muestran en las Figuras 3.17 y 3.18, esta solución no necesita ser instalada en los diferentes equipos donde se utilice, ya que queda persistente en la plataforma de *Google Workspace*, con ellos el usuario podría ejecutar por separado ambos scripts si fuera necesario.

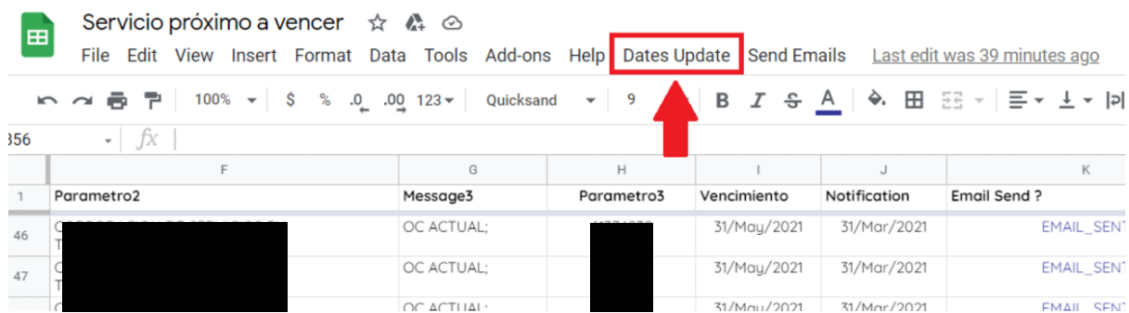


Figura 3.17 Botón para ejecutar script de actualización de fechas.

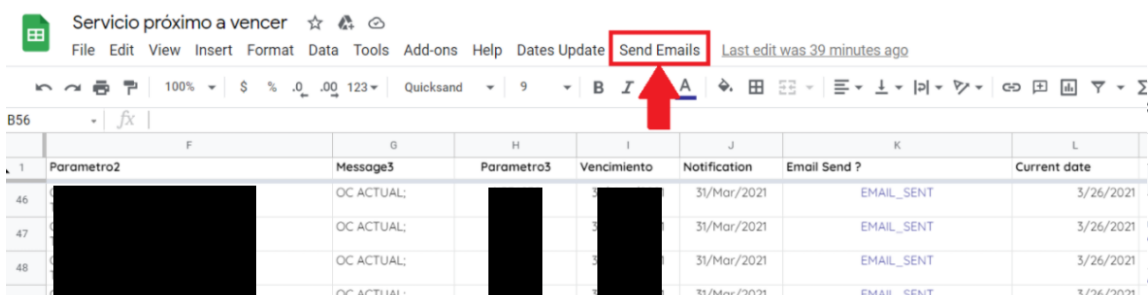


Figura 3.18 Botón para ejecutar script de actualización de fechas.

A continuación se muestra un ejemplo de cómo se vería un correo al ser enviado mediante el script.

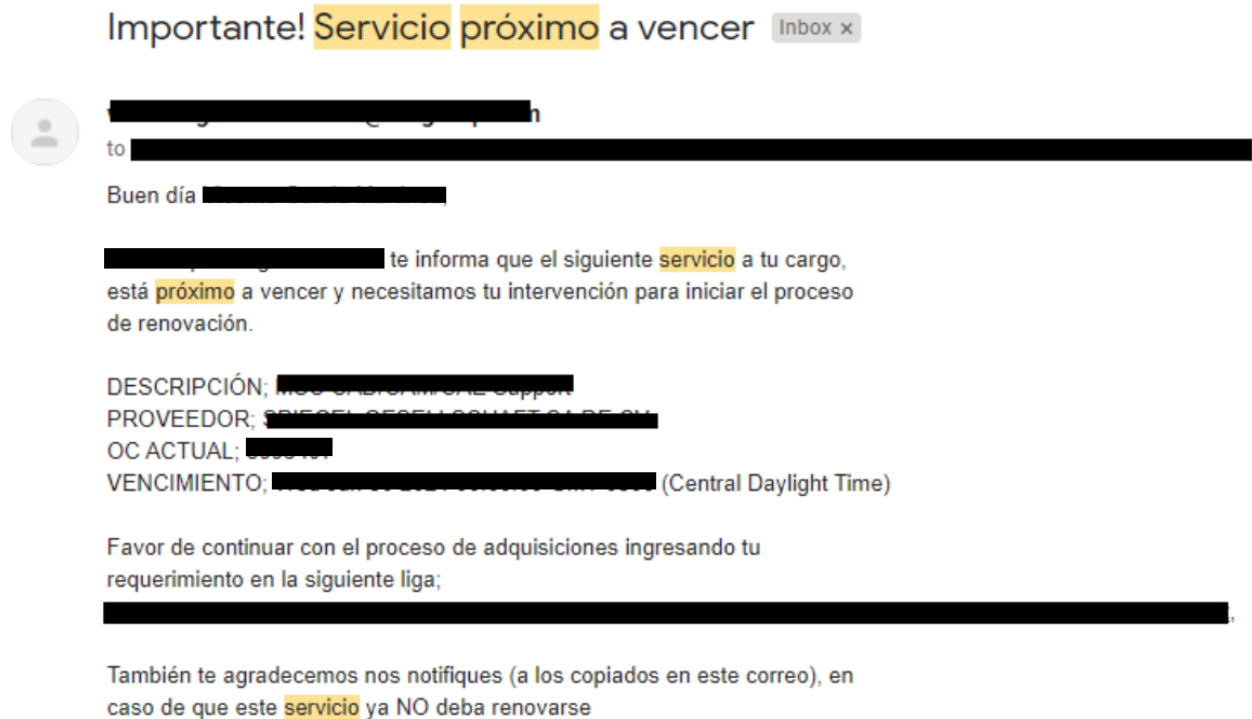


Figura 3.19 Correo resultado. Se ocultaron datos confidenciales.

4. Reportes en tiempo real sobre status de proyectos

Como en toda empresa, se llevan a cabo múltiples proyectos al mismo tiempo, sin embargo, para los gerentes de proyecto, líderes de equipo e incluso para cualquier integrante de equipo es de suma importancia saber la información más actual en relación con cualquier tema que sea necesario para la continuación de cada proyecto.

Actualmente, todos los datos de cómo están distribuidas las aplicaciones, sus conexiones, repositorios de código, archivos estáticos e incluso en qué mejora se encuentran y cuantos issues tienen sin resolver, se encuentran en archivos en la nube, específicamente Google Sheets, como se muestra en la Figura 4.1.

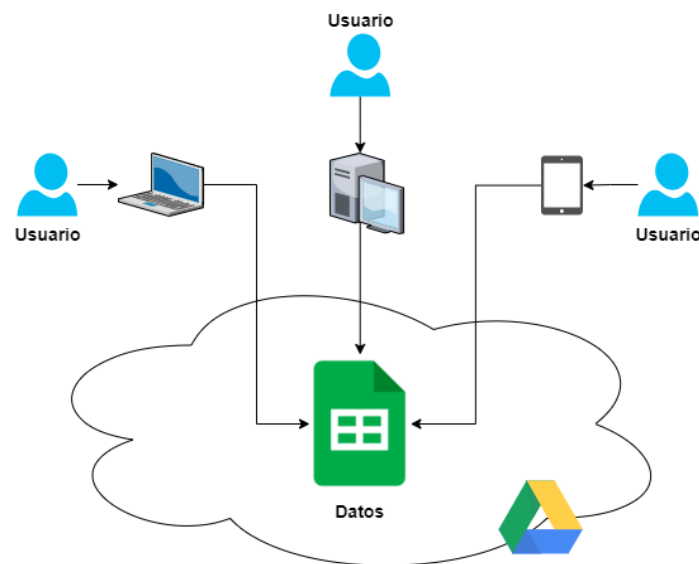


Figura 4.1 Interacción del usuario con los datos.

Esta información fue constantemente revisada y actualizada por los usuarios mencionados anteriormente, sin embargo, el principal problema que se tenía era que todos los datos se encontraban organizados en forma de tablas, con lo cual dificultaba hallar de manera rápida y eficaz información en específico y esto para la toma de decisiones no permitía tener un panorama realista de la situación de los diferentes proyectos.

La empresa ya comenzaba a experimentar con los diferentes tipos de reportes que se hacían de esta información. No obstante, dichos reportes tenía algunos defectos; el más

significativo era que tenían información estática, esto causaba que a lo largo del tiempo no se iban actualizando y generaban un retraso en la información con la que se contaba a la mano. A su vez estos eran también generados en la plataforma de Google Sheets, como se muestra en la Figura 4.2, con lo cual en un solo archivo se contenía demasiada información y era propensa a confusiones.

El objetivo del proyecto fue generar reportes dinámicos y automáticos, con lo cual se iba a garantizar que la información mostrada en ellos fueran la más actual sin tener la necesidad de volver a crear otro reporte. A su vez, se solicitó la creación de un archivo donde se ubicarían todos los reportes por separado con lo cual, a través de un simple clic se llevaría de una manera automática al reporte correspondiente; de esta manera se tendría una mejor organización de la información sin tener que saturar un archivo en específico.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
	Nombre	AMS ID	ABC Number	AMS - System Name	Application Description	Availability	Type	Response	Users	Business Area	Test Environment URL		Prod Environment URL		GitHub Re
											Intro URL TEST	Extra URL TEST	Intro URL PROD	Extra URL PROD	
4		5365			Business Enabling	-	App	Yes	Corporate						
5		5366			Business Enabling	-	App	Yes	Corporate						
7		5367			Business Enabling	-	App	Yes	Corporate						
13		6833			Business Enabling	-	App	Yes	Plants						
14		6855			Business Enabling	-	App	Yes	Corporate						
16		6841			Business Enabling	-	App	Yes	Corporate						
19		9104			Business Enabling	-	App	Yes	Corporate						
21		9270			Business Enabling	-	App	Yes	Corporate						
24		9278			Business Enabling	-	App	No	Plants						
25		9293			Business Enabling	-	App	Yes	Corporate						
26		9293			Business Enabling	-	App	Yes	Corporate						

Figura 4.2 Ejemplo de los datos. Se ocultaron datos confidenciales.

4.1 Entendimiento de requerimientos

El primer paso para el desarrollo de este proyecto, se dio mediante videoconferencias en Google Meet. En ellas se dió a conocer el detalle del estado actual de los datos y sus ubicaciones, así como los objetivos a los que se quería llegar.

Se explicó que actualmente no contaban con reportes dinámicos, por lo cual tenían que estar creando constantemente nuevos para así tener la información más actual, sin embargo, ya contaban con archivos fuente, es decir, esos archivos eran los que actualizaban para posteriormente crear los reportes, esto fue fundamental pues ya se contaba con la información actual.

Para cada uno de los archivos se requería generar un reporte dinámico mediante el cual los usuarios, gerentes o cualquier persona interesada pudiera consultar la información y poder interactuar con ella, aplicando filtros, selecciones, entre otros, para así lograr no solo poder saber el panorama general sino también información en específico.

Se solicitaron dos entregables principales:

1. Un archivo Sheets donde se almacenaran todos los reportes y así lograr un fácil acceso, ya sea mediante el uso de hojas o links con el redireccionamiento correspondiente, de esta manera se podría compartir el archivo con los diferentes usuarios. A su vez este archivo debería contener una descripción de cada análisis para poder saber de manera rápida el funcionamiento de cada uno.
2. Un reporte dinámico para poder interactuar la información mediante el uso de filtros, selecciones, entre otros. Este análisis tenía que ser sencillo de leer con uso de diferentes tipos de gráficas y tablas, con la condición de que estuviera conectado directamente a los archivos fuentes para garantizar que la información estaría al día.

4.2 Comprensión y obtención de los datos

Para esta fase se me proporcionó la lista y el acceso a los archivos fuentes de los cuales se tuvieron que realizar los análisis y reportes necesarios, cada uno de ellos contenía información desde 2018 sobre diferentes proyectos y tecnologías. La lista de los archivos fuentes fue la siguiente:

- Listado_WPO_Aplicaciones_Web_en_Produccion: corresponde al archivo que contiene la información de las diferentes apps y batch que se tienen, así como la distribución de las diferentes mejoras y sus pruebas realizadas.
- Data Issues: este archivo contiene la información de los diferentes problemas que han salido en las liberaciones de procesos batch o aplicaciones Web.
- Mexico_Applications_to_GitHub (local): en este documento se almacena la información de los nombres y estatus de los repositorios de aplicaciones y procesos batch, tanto locales como en la plataforma de DealerConnect.
- Tracking Status PAR WPO MX: en este Google Sheets se tiene la información de todos los tickets tipo PAR que se solicitan para los diferentes proyectos.
- Listado de Librerías Compartidas Commons & Resources: en este archivo se tiene la información de las librerías y recursos compartidos que utilizan las aplicaciones y cómo es que están almacenadas en los diferentes servidores.
- Issue_Log_ProyectosWeb_TMO MX/RUN: cuando se realizan mejoras ocurren issues en las aplicaciones, algunos se atienden de inmediato, pero otros pasan a este archivo para ser atendidos en futuras mejoras; se encuentra dividido entre las diferentes áreas de programación.
- Pruebas Funcionales - Web Local Java 2021: para un control de las pruebas en este archivo se almacena la información de a quién están asignadas las pruebas funcionales, de carga, memoria y estrés, además del estatus de cada uno de ellas y a qué proyecto corresponden.

Adicionalmente, se utilizó el lenguaje de programación llamado Python para la obtención y exploración de información mediante el uso de *web scraping*. Esto me permitió obtener, en tiempo real, información técnica, de estructura y tecnologías que están utilizando las aplicaciones web y procesos batch de la empresa, revisando ambientes tanto productivos como en ambientes de prueba.

Se utilizó *web scraping*, ya que es una herramienta enfocada para la obtención y extracción de datos de un entorno web con ventaja que se realiza de una manera automatizada (Castrillo Fernández, 2015), esto me ayudó a tener los datos estructurados en la forma en que se necesitaban además de obtener los más actualizados, ya que solo se necesita ejecutar el código y en minutos obtendremos toda la información.

4.3 Preprocesamiento de los datos

Para tener una mejor organización de todos los archivos proporcionados anteriormente se optó por reunirlos en un solo archivo Google sheets, llamado “Analysis”, se muestra en la Figura 4.3, cada uno de las fuentes principales era copiado a través de funciones *import* con lo cual se garantizó tener la información actual, no obstante para el archivo de los “Data Issues”, se muestra en la Figura 4.4, después de un análisis visual se optó por dejarlo fuera, ya que desde su origen mantenía una estructura idónea para los siguientes análisis.

1	Name	AMS ID	eDC Number	AMS - System Name	Application Description	Availability	Type	Responsive	Users	Business Area	Test - Intro URL	Test - Extra URL	Prod - Intro URL	Prod - Extra URL	GitHub Repo Name	Repo GitHub URL	Repo GitHub URL - Estado	Repo GitHub URL - Propietario
2							App	No										
3							App	Yes										
4							App	Yes										
5							Batch	N/A										
6							App	Yes										
7							Batch	N/A										
8							App	Yes										
9							Batch	N/A										
10							Batch	N/A										
11							App	No										
12							App	Yes										
13							App	Yes										
14							App	Yes										
15							App	Yes										
16							Batch	N/A										
17							App	No										
18							App	Yes										
19							Batch	N/A										
20							App	Yes										
21							Batch	N/A										
22							App	No										
23							App	No										
24							App	Yes										
25							App	Yes										
26							App	Yes										
27							Batch	N/A										
28							Batch	N/A										
29							App	Yes										
30							App	No										
31							App	Yes										
32							App	No										

Figura 4.3 Archivo “Analysis”.

1	#	Proyecto	Aplicación	Tipo de Issue	Fecha Issue	Fecha Solución	Status	Solucionado por	Descripción error
2	1			Control					No se
3	2								La q
4	3			edados					La m
5	4			PT					Dato
6	5			edados					Ha
7	6			306					No se
8	7			G					No
9	8			ts &					x
10	9			ns					Se p
11	10			ox					Se p
12	11			ed					E

Figura 4.4 Archivo “Data Issues”. Se ocultaron datos confidenciales.

Mediante la función de Google Sheets llamada *importrange*, se muestra en la Figura 4.5, se nos permite seleccionar una hoja de cálculo en específico y extraer los datos que nosotros deseamos especificando las celdas que nosotros queramos obtener. Mediante esta función logré tener en un solo archivo toda la información de las diferentes fuentes proporcionadas; esta información está actualizada en tiempo real, con lo cual los análisis se logran hacer con lo más reciente.

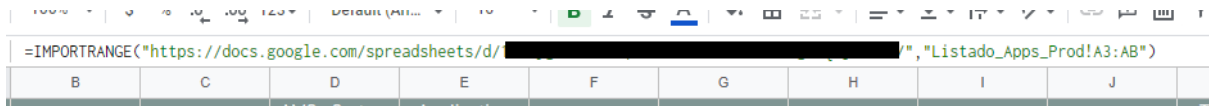


Figura 4.5 Ejemplo de la función IMPORTRANGE usada. Se ocultaron datos confidenciales.

4.4 Análisis

Parte fundamental de los requisitos era que los análisis y reportes puedan ser visuales lo cual ayudaría a los usuarios entender la información de manera rápida y precisa por lo cual se estudiaron dos propuestas para la realización de dichos reportes.

La primera propuesta era realizar los reportes en Google Sheets, con lo cual se mantendría el mismo entorno y una conexión directa con el archivo de “Analysis”, sin embargo, no se siguió con esta propuesta pues me percaté que al desarrollar una interfaz gráfica en ese documento se dificulta la navegación e interacción entre usuario y reporte (Figura 4.6).

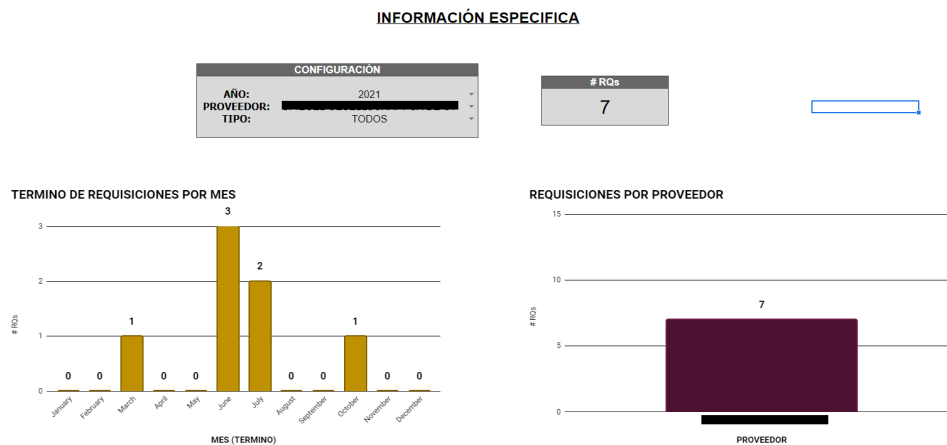


Figura 4.6 Ejemplo de reporte realizado en Sheets. Se ocultaron datos confidenciales.

La segunda opción era realizar los *dashboard* o reportes en la plataforma de Google Data Studio, se muestra en la Figura 4.7, con esto también se garantizó una conexión directa entre el *dashboard* y el archivo fuente de datos, además de poder al guardar todos los *dashboards* en una sola ubicación, esto mantendría controlado los accesos y modificaciones a los mismos.

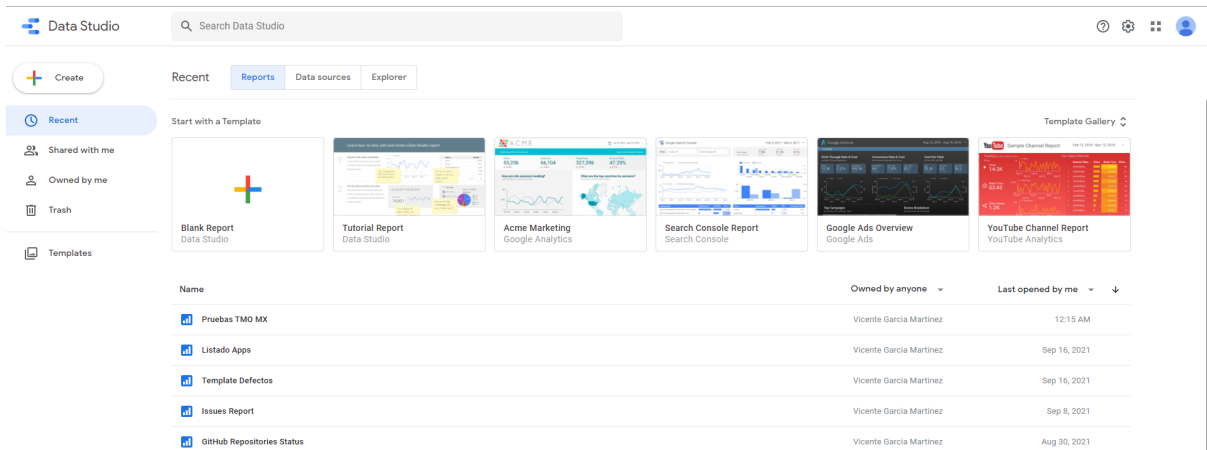


Figura 4.7 Plataforma de Google Data Studio.

Antes de comenzar a desarrollar los *dashboard* en esta plataforma primero se verificó las restricciones que tenía la empresa sobre esta tecnología, al no tener restricciones se decidió continuar el proyecto utilizando esta herramienta, posteriormente se revisó qué tipo de fuentes se podrían utilizar y de qué manera las iba a consumir la interfaz, con esto se pudo conectar de manera directa al archivo fuente, pero también podrían utilizarse datos de otros orígenes ya sean archivos estáticos o incluso otros servicios de Google como Google Analytics. Las conexiones se muestran en la Figura 4.8.

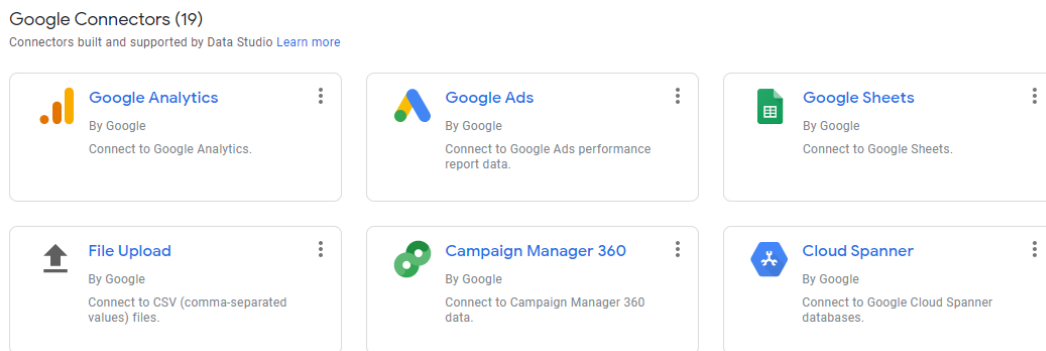
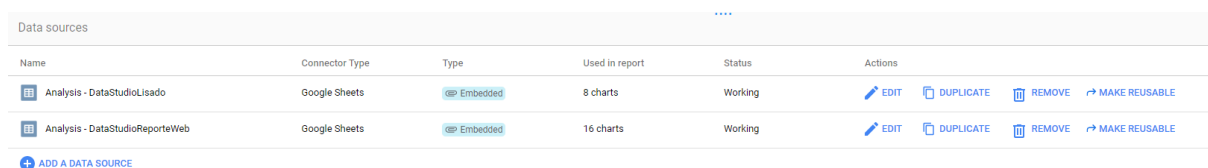


Figura 4.8 Tipo de conexiones.

4.5 Obtención de resultados

Para esta fase se comenzó con la creación de cada uno de los reportes en la plataforma de Google Data Studio. El primer paso fue establecer las conexiones entre la fuente de datos “Analysis”, se muestra en la Figura 4.9, así como los rangos a utilizar de cada uno de ellos, al ser un entorno de Google esta conexión fue directa y sencilla.



Name	Connector Type	Type	Used in report	Status	Actions
Analysis - DataStudioListado	Google Sheets	Embedded	8 charts	Working	EDIT DUPLICATE REMOVE MAKE REUSABLE
Analysis - DataStudioReporteWeb	Google Sheets	Embedded	16 charts	Working	EDIT DUPLICATE REMOVE MAKE REUSABLE

[+ ADD A DATA SOURCE](#)

Figura 4.9 Conexión con el archivo “Analysis”.

Posteriormente, se comenzó con el desarrollo de cada uno de los *dashboard* buscando darles un tema de color específico para cada uno de ellos, así como enfocarse en la creación y visualización con diferentes gráficas, marcadores y paneles de control para la interacción entre el usuario y el reporte. Un ejemplo se muestra en la Figura 4.10.

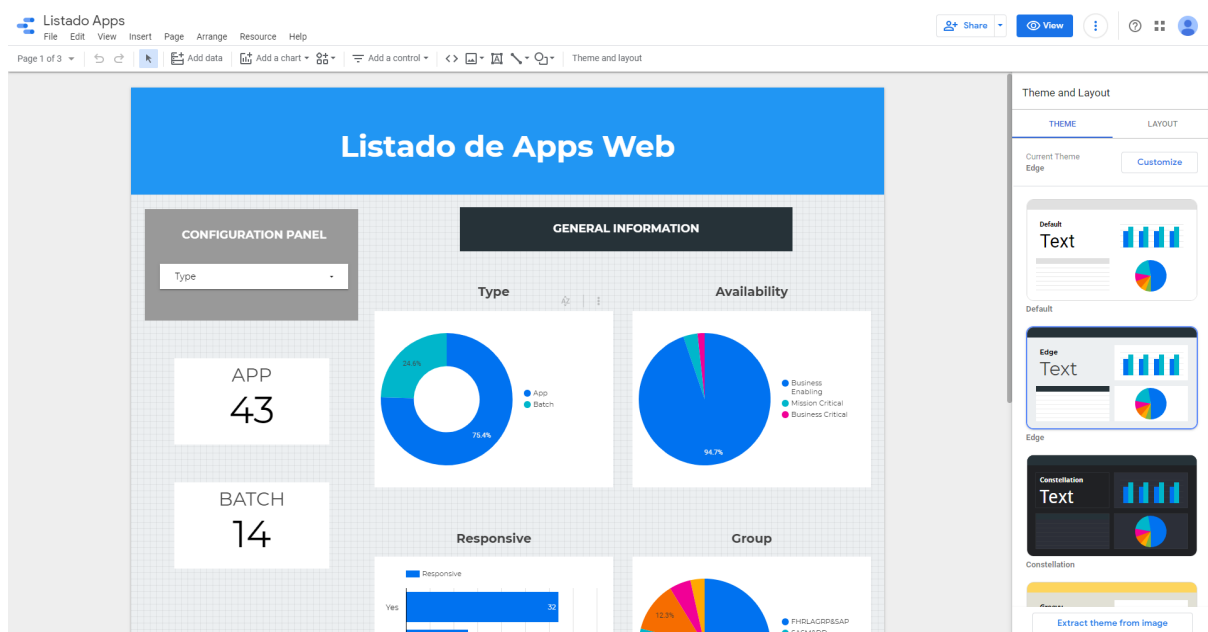


Figura 4.10 *Dashboard* para Listado de App Web.

Además, de poder crear diferentes *dashboard*, también se utilizó la parte de crear secciones dentro de un mismo archivo. Con esto se logró separar la información de una manera más clara para una mejor visualización y no tener que saturar una página con toda la información, como se muestra en la Figura 4.11.

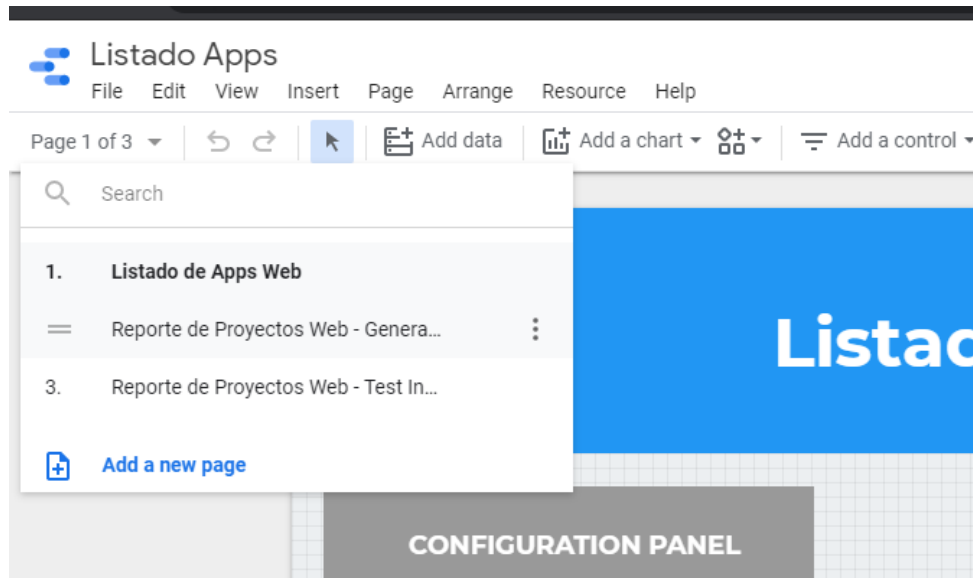


Figura 4.11 Secciones en *dashboard* de “Listado de App Web”.

Otra característica importante fue la creación de un panel de configuración, en el cual a través de selecciones se puede especificar la información que se desea filtrar a través de área, año, plataforma, y otros, y así unos análisis específicos como se muestra en la Figura 4.12.

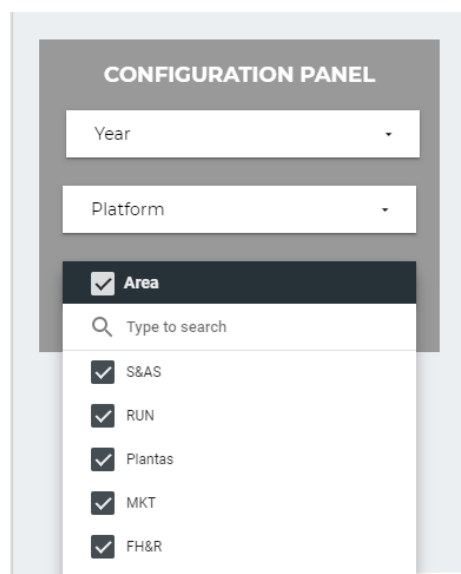


Figura 4.12 Ejemplo de panel de control implementados.

En el caso de *DataStudio* se utiliza una interfaz mucho más interactiva para realizar cálculos, gráficas, etc, pero en todos los casos se siguieron los mismo pasos,

1. Se seleccionó el gráfico deseado, figura 4.13, esto depende de lo que se quiera mostrar, ejemplos como *scorecard*, gráficas de pastel, de barras, geográficas o incluso tablas.

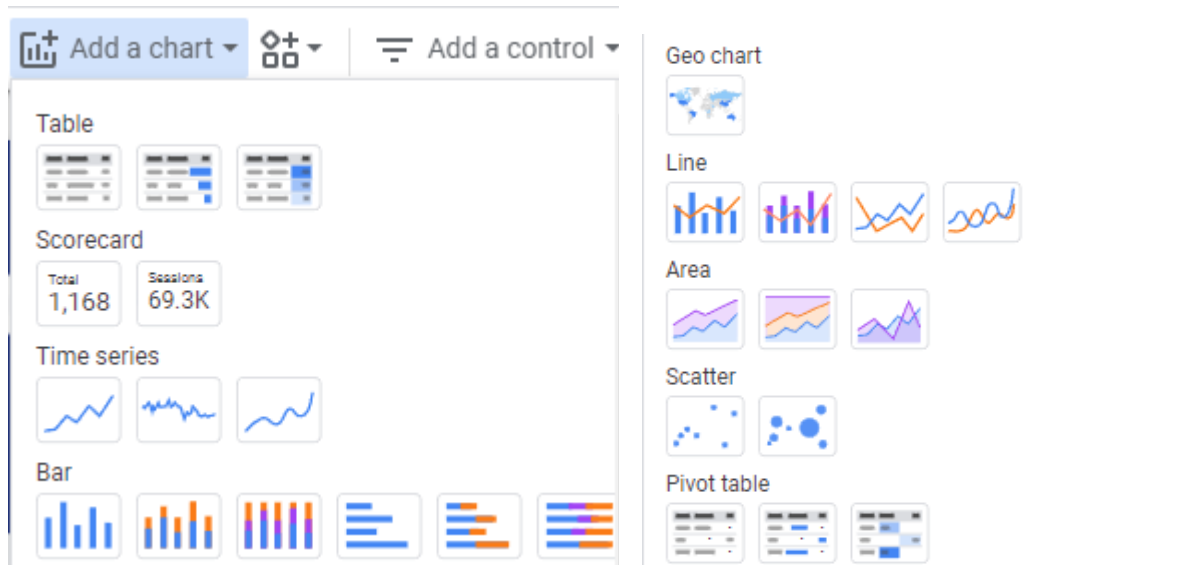


Figura 4.13 Ejemplo de gráficos a utilizar.

2. Ya con el gráfico seleccionado se tuvo que enlazar con la fuente de datos, previamente cargada y con ello se nos mostraba las diferentes columnas de la misma, figura 4.14.

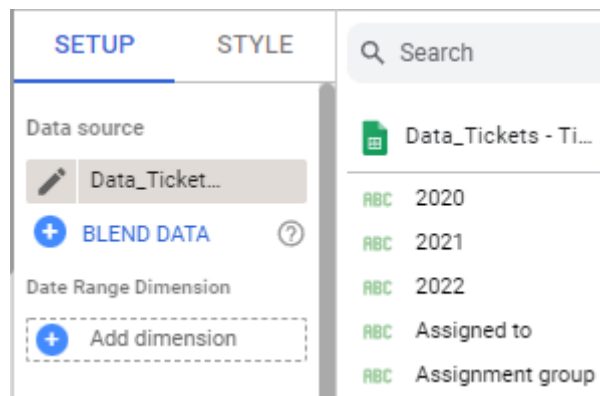


Figura 4.14 Ejemplo de selecciones de fuente de datos (izquierda) y columnas de la misma (derecha).

3. Ya con las columnas seleccionadas se habilitaba un menú en donde se ponían las diferentes operaciones a realizar con los datos, figura 4.15, desde conteos, sumas, cálculo de máximos, mínimos, etc. Para este paso fue necesario determinar el tipo de dato final así como sus diferentes cálculos o comparaciones.

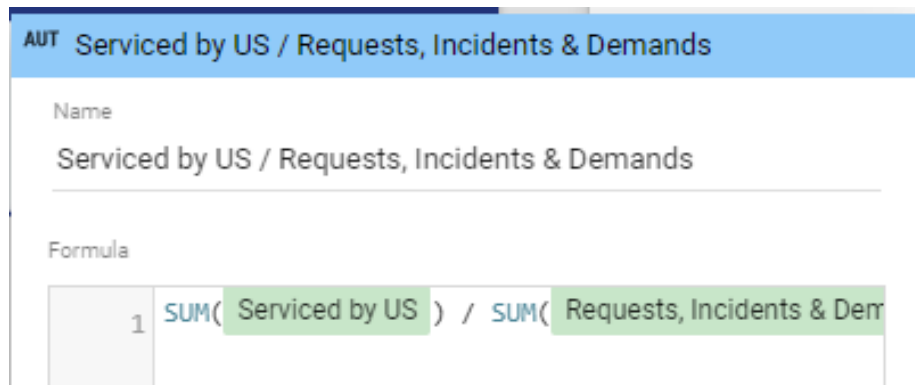


Figura 4.15 Menú para la configuración de los cálculos implementados.

4. *DataStudio* cuenta con la posibilidad de crear diferentes filtros, por lo que dependiendo el objetivo de la gráfica fue necesario crear diferentes filtros, esto se lograba mediante un menú, figura 4.16, en el que se creaban los filtros aplicando diferentes comparaciones.

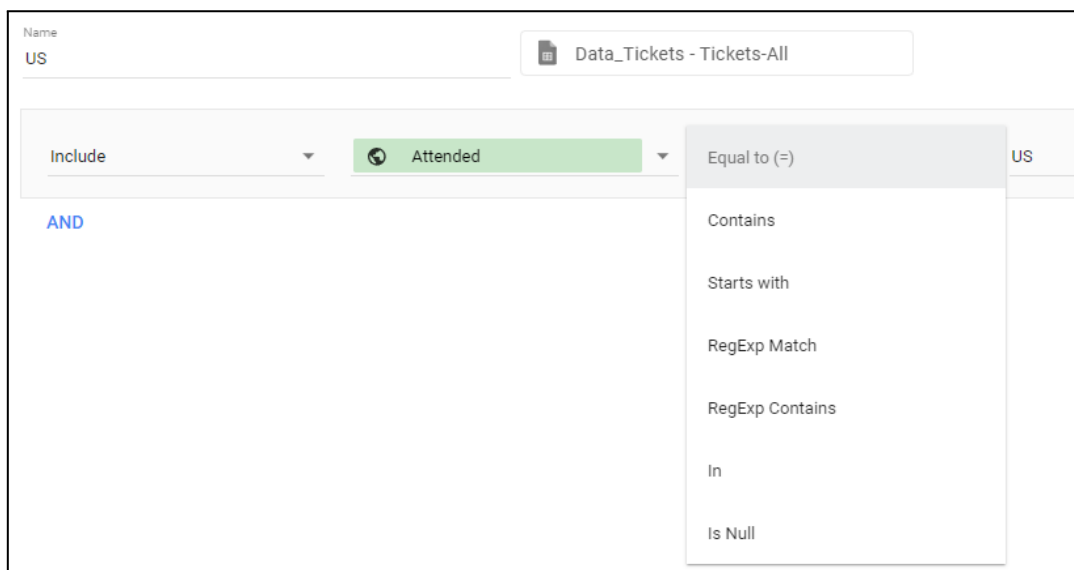


Figura 4.16 Menú para la creación de filtros.

5. En algunos se tuvo el desafío de que era necesario unir diferentes fuentes de datos, para esto se utilizó la herramienta de *Blend Data*, figura 4.17, esta función permite combinar datos con la condición de que al menos deben tener una columna con el mismo nombre y tipo de datos, adicionalmente se selecciona el tipo de join a utilizar.

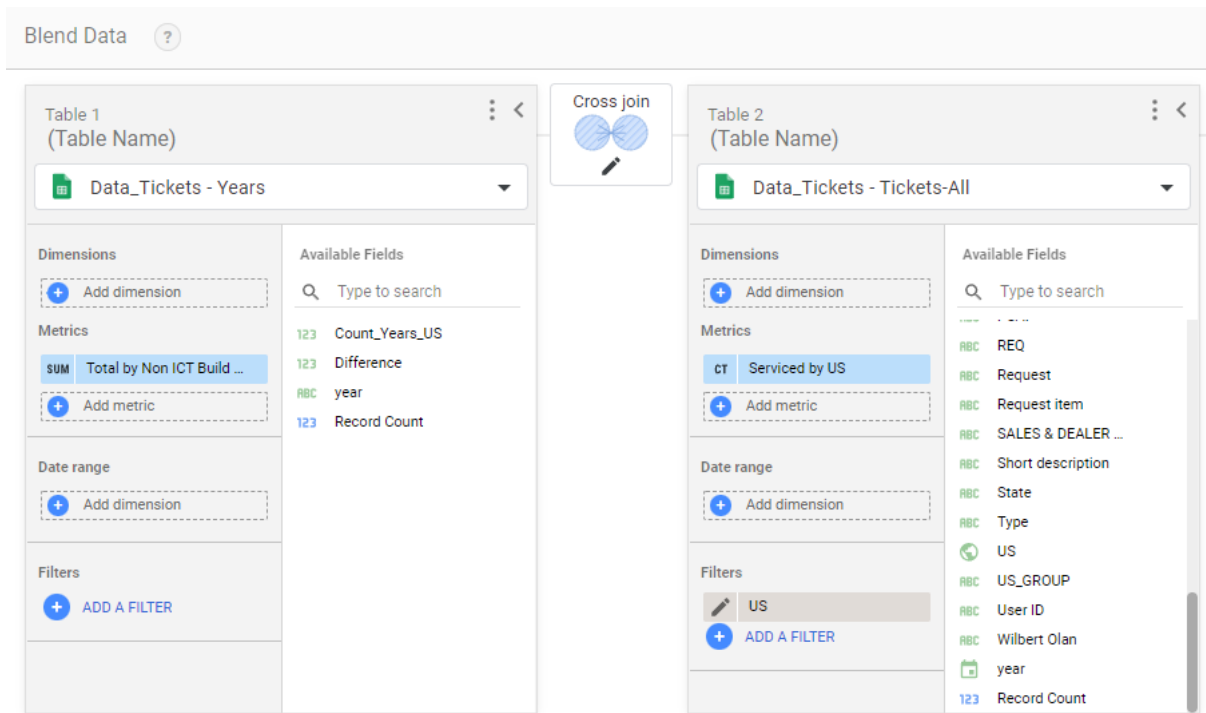


Figura 4.17 Menú de *Blend Data* en *DataStudio*.

4.6 Despliegue

Para la distribución de los *dashboard* generados se encontraron cuatro distintas maneras de hacerlo, cada una de ellas con sus ventajas y restricciones, como se muestra en la Figura 4.18.

1. La primera opción consistía únicamente en compartir el reporte a través de un link, con lo cual cualquier persona que lo contenga podrá tener acceso visual a dichos reportes.
2. La segunda opción era descargarlo en formato PDF y poderlo compartir de esta manera, sin embargo, esta acción restringía la posibilidad de que fuera un reporte dinámico y en tiempo real.
3. La tercera opción era configuración de envío automático de reporte, de esta manera se puede configurar qué día y a qué hora enviar el reporte a través de correos deseados.
4. Y por último era invitar a los distintos usuarios mediante sus correos de Gmail de la empresa para tener acceso a los reportes, este acceso se podría configurar con los permisos que nosotros quisiéramos otorgarles.

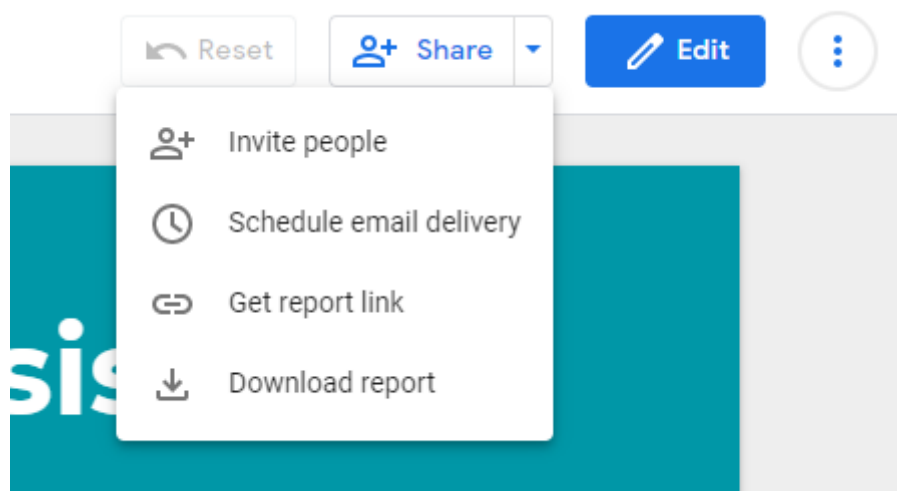


Figura 4.18 Opciones para compartir los reportes.

Para compartir los reportes creados se optó por la cuarta opción ya que ésta nos permitía configurar los distintos permisos que tendría cada usuario y así poder restringir u otorgar, dependiendo sea el caso, como se muestra en la Figura 4.19. Adicionalmente, esta opción también nos proporcionó un link, el cual únicamente tiene acceso visual y ese lo pudimos compartir a las distintas personas que nosotros queramos que también vean los análisis sin tener que agregarlos directamente con permisos superiores.

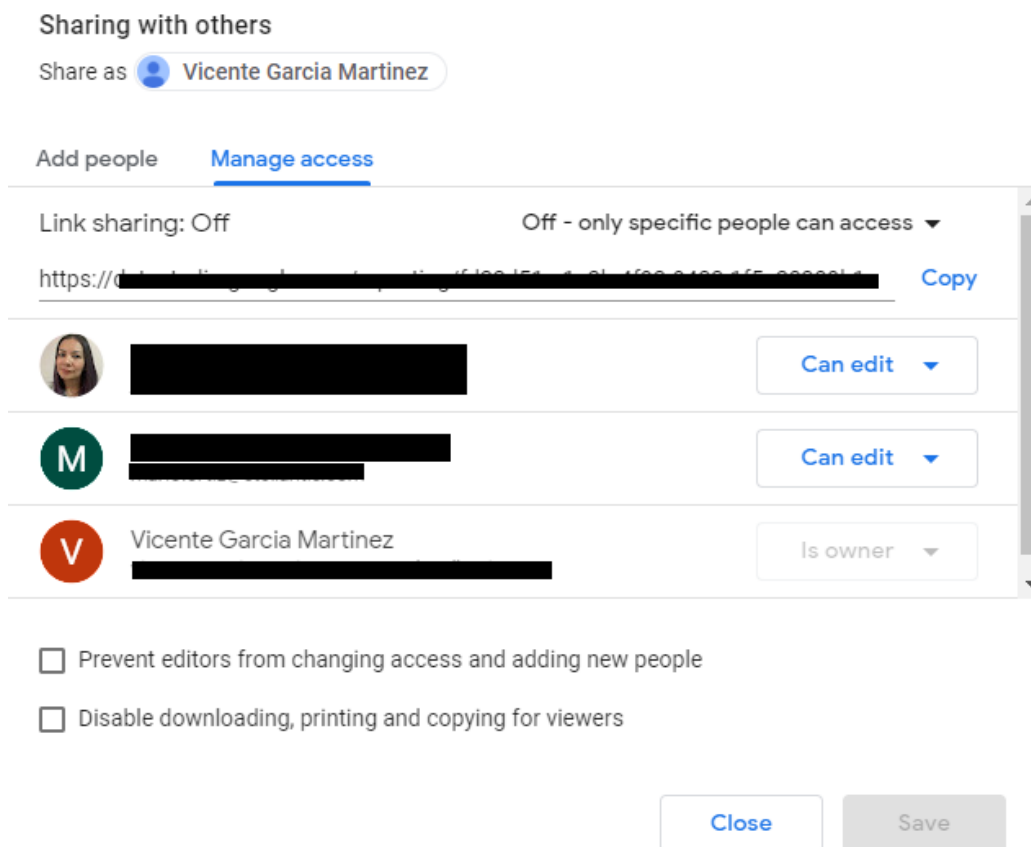


Figura 4.19 Opciones para compartir los reportes mediante control de permisos. Se ocultaron datos confidenciales.

5. Resultados

En este capítulo se explica la forma en la que se entregaron ambos proyectos, tanto de automatización de procesos de notificaciones, así como el de reportes en tiempo real. Además, se muestra como es que la empresa los están utilizando durante el tiempo que ha transcurrido de dichos proyectos.

5.1 Automatización de proceso de notificaciones

Para la entrega del proyecto de automatización de procesos de notificaciones, se procedió a la creación de una guía de uso, Figura 5.1, con la cual cualquier persona que tuviera los accesos correspondientes, a los archivos de este proyecto, podía entender el funcionamiento a nivel alto y el cómo poder implementarlo.

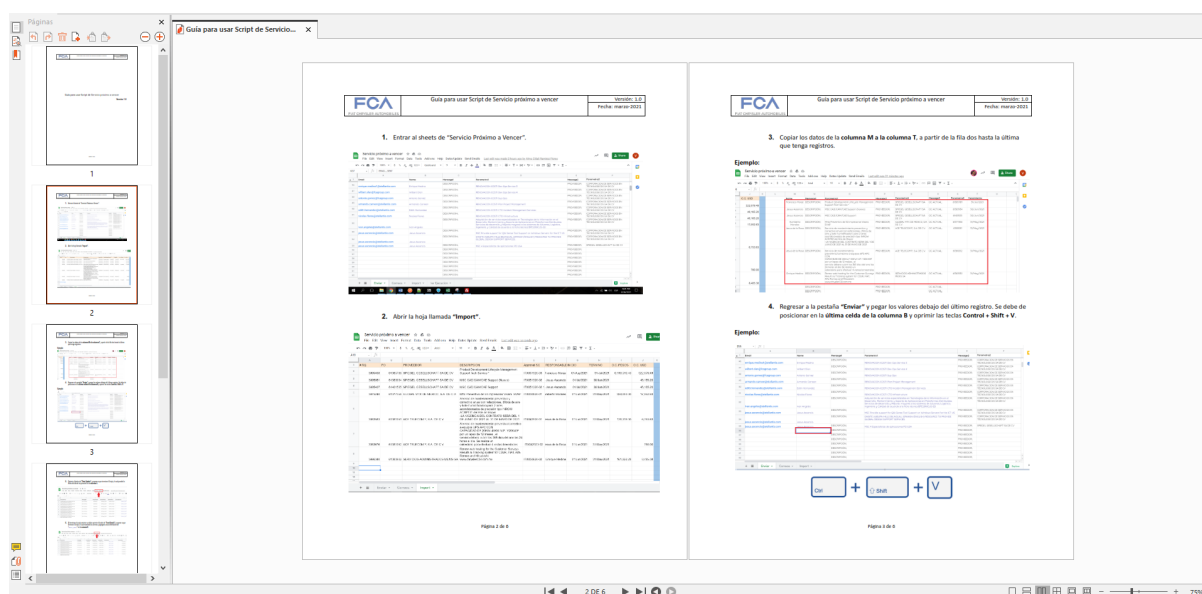


Figura 5.1 Guía de Uso del proyecto de Automatización de proceso de notificaciones.

En esa guía se explica el funcionamiento y flujo que se lleva a cabo durante la ejecución de la automatización, es decir, desde la obtención de los datos dependiendo los

criterios solicitados la generación y envío de los correos. El primer archivo entregado fue un *dashboard*, Figura 5.2.

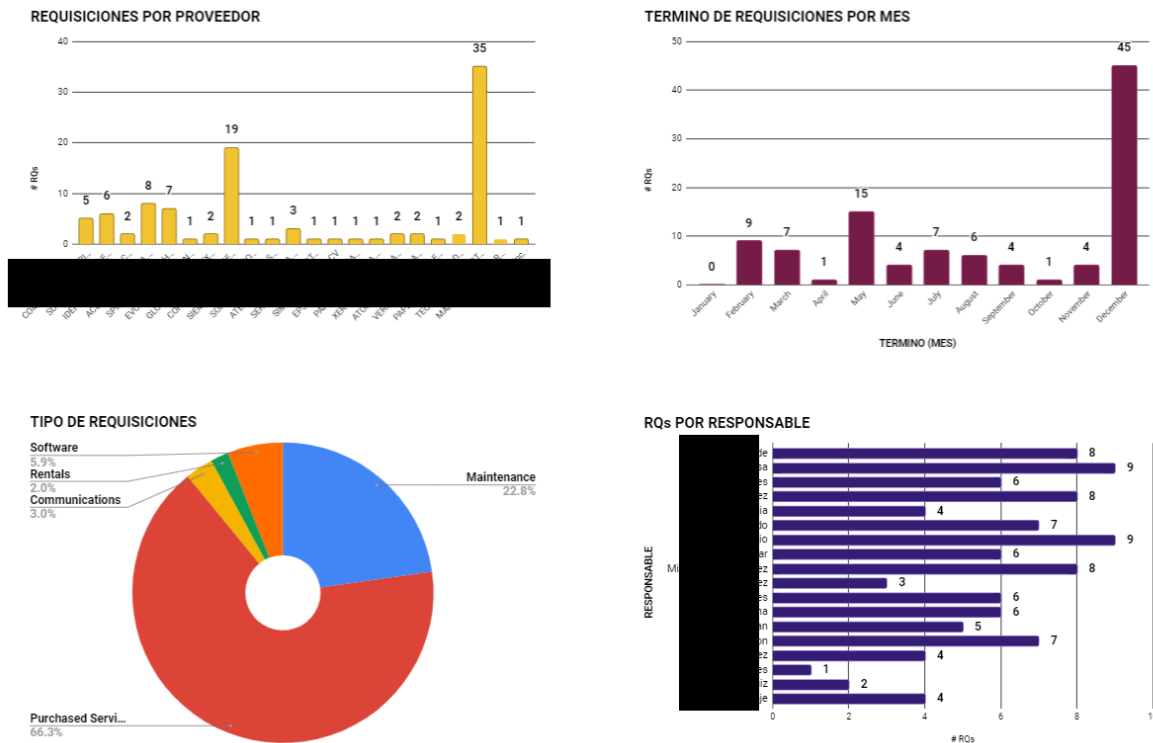


Figura 5.2 *Dashboard* de Requisiciones General. Se ocultaron datos confidenciales.

Mediante este reporte se puede ver cómo se estaban comportando las requisiciones dependiendo de los proveedores, cantidad y tipo de requisiciones, fecha de inicio y fecha final, entre otros aspectos importantes, los cuales permitían tener un panorama más específico de la situación no solo del año en curso sino de años posteriores y años futuros en específico desde 2019 hasta 2024 .

El siguiente archivo que se entregó fue un archivo de Google Sheets llamado “Servicio próximo a vencer”.

Name	Message1	Parametro1	Message2	Parametro2	Message3	Parametro3	Vencimiento
[REDACTED]	DESCRIPCIÓN; [REDACTED]	[REDACTED]	PROVEEDOR; [REDACTED]	[REDACTED]	S.A. OC ACTUAL; [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	DESCRIPCIÓN; [REDACTED]	[REDACTED]	PROVEEDOR; [REDACTED]	[REDACTED]	DE OC ACTUAL; [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	DESCRIPCIÓN; [REDACTED]	[REDACTED]	PROVEEDOR; [REDACTED]	[REDACTED]	OC ACTUAL; [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	DESCRIPCIÓN; [REDACTED]	[REDACTED]	PROVEEDOR; [REDACTED]	[REDACTED]	OC ACTUAL; [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	DESCRIPCIÓN; [REDACTED]	[REDACTED]	PROVEEDOR; [REDACTED]	[REDACTED]	NG OC ACTUAL; [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	DESCRIPCIÓN; [REDACTED]	[REDACTED]	PROVEEDOR; [REDACTED]	[REDACTED]	A OC ACTUAL; [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	DESCRIPCIÓN; [REDACTED]	[REDACTED]	PROVEEDOR; [REDACTED]	[REDACTED]	OC ACTUAL; [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	DESCRIPCIÓN; [REDACTED]	[REDACTED]	PROVEEDOR; [REDACTED]	[REDACTED]	OC ACTUAL; [REDACTED]	[REDACTED]	[REDACTED]

Figura 5.3 Archivo “Servicio próximo a vencer”. Se ocultaron datos confidenciales.

En este archivo se hacía la obtención y filtrado de los datos finales los cuales servirían como entrada para los dos programas posteriores. Aquí se definía la estructura que tendrían así como los distintos campos que se deberían obtener ya que son los necesarios para la automatización final de los correos. Con respecto al filtrado en relación con los criterios pedidos por el cliente.

Por último, en los archivos entregados se encuentran los dos códigos programados mediante Google Script, Figura 5.4, mediante los cuales se realiza la actualización de campos y la generación de los correos que posteriormente son mandados de forma automática.



Figura 5.4 Nombre de los dos códigos entregados.

Ambos códigos se encuentran totalmente documentados y comentados, con esto se garantiza que se podrán utilizar y actualizar en futuras mejoras sin tener que reescribir todo y logrando hacer que se entienda su flujo de manera clara y sencilla. El código llamado Date.gs es el encargado de preparar y actualizar los campos necesarios para el proceso, como fechas y operaciones realizadas con los mismos. El segundo código con el nombre Code.gs es el programa principal, mediante este se realiza la configuración y estructura de los correos para posteriormente ser enviados de manera automática, los correos destinatarios dependen del archivo fuente, sin embargo, los correos a los cuales va copiado son estáticos, el cuerpo del

correo así como el asunto son totalmente personalizados dependiendo la fuente y requisición correspondiente.

Para la entrega de todos estos archivos se tuvo una última reunión con los clientes, en donde se les pudo explicar archivo por archivo y se solucionaron y aclararon las dudas correspondientes que llegaron a tener además de darles los accesos correspondientes, adicionalmente se estuvo de forma presencial en la primera ejecución de toda la automatización y se verificó el correcto funcionamiento de la misma, se generaron las evidencias correspondientes para poder dar cierre al proyecto, esto se llevó mediante la plataforma de Google Meet mediante la cual se hizo uso de la función de compartir la pantalla tanto para la visualización de los archivos como de la ejecución de los programas.

5.2 Reportes en tiempo real sobre status de proyectos

Para la entrega del proyecto de reportes en tiempo real sobre status de proyectos se procedió a la creación de un *landing page*, con el cual cualquier persona que tuviera los accesos correspondientes a los reportes de este proyecto podía acceder y visualizarlos, así como poder interactuar y ver los distintos comportamientos.



Figura 5.5 Landing Page.

El segundo archivo entregado fue un Google Sheets llamado “Analysis”, mediante el cual se hacía la obtención de los datos de las distintas fuentes que se tenían, así como su procesamiento para darles homogeneidad y estructura, esto sirvió para que el archivo fuera la entrada del programa de Google Data Studio y poder generar los reportes.

1	Name	AMS ID	eBC Number	AMS System Name	Application Description	Availability	Type	Response	Users	Business Area	Test - Intro URL	Test - Extra URL	Prod - Intro URL	Prod - Extra URL	GitHub Repo Name	Repo GitHub URL	Repo GitHub URL - Estado	Repo GitHub URL Properties
2	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	No	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
3	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
4	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
5	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
6	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
7	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
8	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
9	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
10	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
11	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	No	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
12	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
13	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
14	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
15	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
16	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
17	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	No	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
18	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
19	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
20	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
21	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
22	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	No	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
23	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	No	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
24	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
25	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
26	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	App	Yes	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
27	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
28	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	Batch	N/A	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
29	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]

Figura 5.6 Archivo “Analysis”. Se ocultaron datos confidenciales.

Este archivo está relacionado y conectado de manera directa con Data Studio al igual que con las fuentes principales de datos, con esto se garantiza que la información que se tiene es la más reciente y más actualizada y así se pueden obtener los reportes más recientes.

Este archivo fue principalmente compartido con mi supervisor, ya que al ser el archivo fuente para los distintos reportes se debe de tener un control mucho más estricto y restringido para así poder generar los archivos sin tener información errónea.

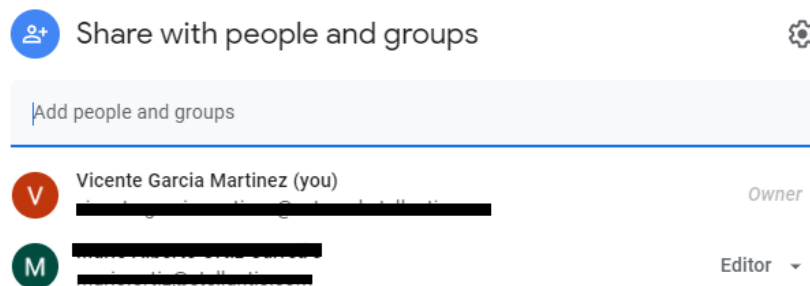


Figura 5.7 Acceso al archivo “Analysis”. Se ocultaron datos confidenciales.

Por último, en la entrega se daba acceso a los diferentes *dashboard* generados en Data Studio en los cuales mediante paneles de control, gráficas de barras, gráficas de pastel, tablas dinámicas e incluso leyendas y contadores se pudo visualizar la situación actual de cada uno de los diferentes análisis y proyectos en curso en la empresa y así poder tomar una mejor decisión con base en datos y estadísticas.

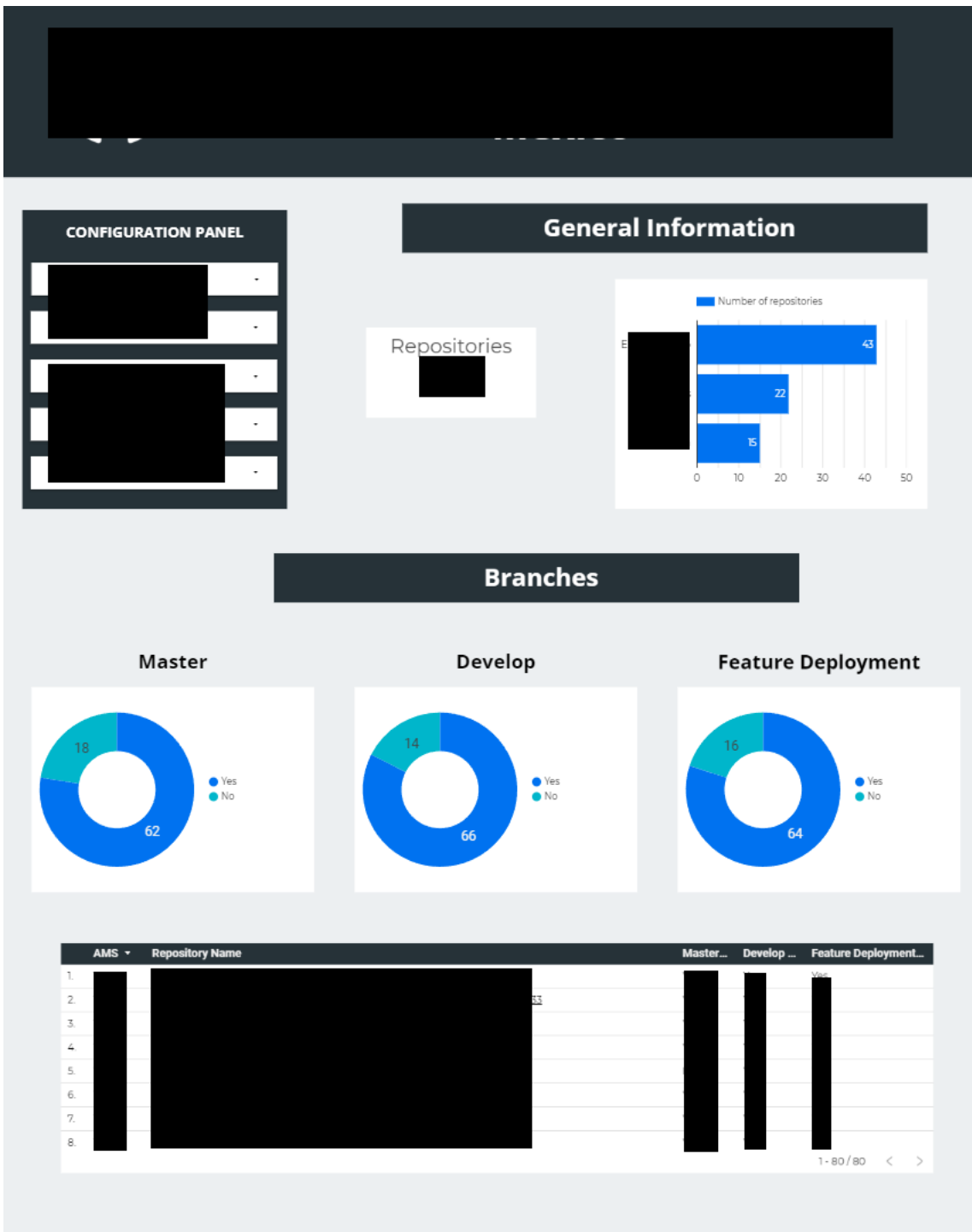


Figura 5.8 Ejemplo *Dashboard* generado. Se ocultaron datos confidenciales.

6. Conclusiones

Durante el desarrollo de estos proyectos y actividades en la empresa me pude percatar que muchos de los procesos pueden ser automatizados para así poder realizar tareas cotidianas en un tiempo mucho menor, evitando así también errores de vista, errores de dedo e incluso retrasos al olvidar realizar una tarea recurrente. En cada uno de los proyectos fue importante conocer la situación actual de los datos y cómo se estaban generando los mismos y posteriormente el conocer la infraestructura que tiene la empresa permitió plantear una solución óptima aprovechando los recursos de la misma.

Por otra parte, un paso fundamental en ambos proyectos fue la generación de documentación evidencia de cada una de las fases ya que esto permitió tener un mayor control sobre las distintas versiones y lo que se iba realizando paso a paso toda esta documentación se mantenía organizada en la nube para así garantizar el acceso a ella en cualquier lugar y tiempo.

Sobre el proyecto de la automatización de procesos y envío automático de correos electrónicos actualmente los clientes han estado completamente satisfechos e incluso posteriormente realizaron mejoras y cambios en los criterios de selección. Adicionalmente, se está pensando en replantear la solución, pero ahora utilizando la infraestructura proporcionada por Microsoft esto debido a que el proceso de automatización les ha servido de mucha ayuda y no lo quieren dejar de usar.

Sobre el uso y aprovechamiento de análisis de datos y su visualización en reportes en tiempo real a través de la plataforma de Data Studio e incluso *dashboard* generados en Google sheets han sido de gran utilidad y actualmente he realizado más de doce proyectos similares, no solo han sido implementados para visualizar la situación actual de distintos proyectos, también para la toma de decisiones a nivel superior, ya que al poder visualizar la información de manera rápida y sencilla a través de gráficas y poder interactuar con los datos y ver su comportamiento se puede tener una mejor visualización de las diferentes situaciones y así poder llegar a una conclusión mucho más óptima y que genere beneficio para la empresa.

Con los análisis de datos se han podido ver cómo trabajan las distintas áreas de desarrollo y así poder detectar cargas de trabajo bajas, medias y altas; con ello se ha podido tomar una mejor decisión a la hora de distribuir los esfuerzos de trabajo, también se ha logrado visualizar cómo se han comportado la detección de problemas y sus correspondientes soluciones. Estos análisis han sido de gran ayuda a la hora de la realización de migraciones a nuevas versiones de código e incluso al uso e implementación de nuevas tecnologías como lo son GitHub, SonarQube entre otros.

Uno de los pasos más importantes en el desarrollo del proceso del análisis de datos es el preprocesamiento de la información pues esto a futuro nos facilita implementarlos y realizar códigos o fórmulas que nos ayuden a obtener las características relevantes o importantes y para ello es importante realizar cada una de las tareas como la limpieza de la información, es decir homologar los datos, detectar nullos, incompletos, entre otros, pero sobre todo entenderlos antes de realizar cualquier modificación o algoritmo.

Por último, una de las características más importantes que he aprendido a lo largo del desarrollo de análisis de datos es que toda empresa cuenta con datos históricos los cuales son fundamentales y de gran ayuda a la hora de realizar dichos análisis, pues brindan información importante de cómo se ha ido comportando la empresa a lo largo del tiempo y con ello lograr predecir sus comportamientos en los futuros años, e incluso detectar situaciones y prepararse para resolverlas a corto o mediano plazo e incluso prevenirlas, todo esto gracias a la información que se obtiene con los datos.

Bibliografía

- Google. (2022). Google Apps Script overview. Google Developers. Recuperado 27 de octubre de 2021, de <https://developers.google.com/apps-script/overview?hl=es-419>
- Google. (2022). Le damos la bienvenida a Data Studio - Ayuda de Data Studio. Le damos la bienvenida a Data Studio. Recuperado 27 de octubre de 2021, de <https://support.google.com/datastudio/answer/6283323?hl=es>
- Hegde, P. (2022, 5 enero). The Importance of Data Analysis in Problem Solving | Towards Data Science. [https://Towardsdatascience.Com/](https://towardsdatascience.com/). Recuperado 28 de octubre de 2021, de <https://towardsdatascience.com/the-importance-of-data-analysis-in-problem-solving-2f8ad041708d>
- Maryville University. (2021, 22 marzo). Top 4 Data Analysis Techniques. Maryville Online. Recuperado 15 de noviembre de 2021, de <https://online.maryville.edu/blog/data-analysis-techniques/>
- Inc, S. & SPSS Inc. (2000). *CRISP -DM 1.0*. SPSS Incorporated.
- Castrillo Fernández, O. (2015, diciembre). WebScraping: Applications and Tools. European Public Sector Information Platform. https://data.europa.eu/sites/default/files/report/2015_web_scraping_applications_and_tools.pdf
- Redacción. (2020, 6 agosto). La importancia del análisis de datos para las empresas. Anáhuac Mayab. <https://merida.anahuac.mx/noticias/aprender-analisis-datos-empresas>
- Jacob Escobar. (2021, 12 octubre). La toma de decisiones y el análisis de datos como la clave del éxito empresarial. EGADE Insights. <https://blog.egade.tec.mx/analisis-de-datos-como-clave-del-exito-empresarial>
- Souza, M. G. (2021, 14 diciembre). What do companies really expect from Data and Analytics? Medium. <https://towardsdatascience.com/what-do-companies-really-expect-from-data-and-analytics-382c3cd07c6b>