

Información general del proyecto

Nombre del grupo: Grupo N°2 Fernandez

Integrantes:

- Carrizo Martina
- Ibañez Emilio
- Ibañez Guillermo
- Ledesma Israel
- Vera Ledesma pablo

Entidad / Institución vinculada:

Área de electricidad, municipalidad de Fernandez

Docente Tutor: Mubarqui Fernando

Año lectivo: 2025

Título del proyecto: Propuesta de base de datos estructurada y digitalizada

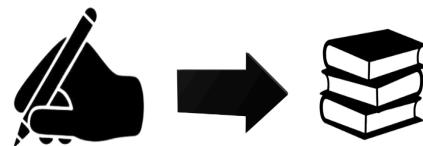
Objetivo general y motivación

El área de electricidad es una entidad dedicada a la reparación y mantenimiento del alumbrado público de la Ciudad de Fernandez y alrededores realizando tareas generales como Cambio de luminarias (focos de 8ow) y cableado. Tiene un equipo trabajando las 24 horas y registran donde se realizaron estas reparaciones y qué insumos utilizan. Esto es relevante para realizar análisis estadísticos dentro de un periodo de tiempo.

El presente proyecto busca brindar una forma más efectiva y rápida de registrar el mantenimiento realizado en la ciudad de Fernandez y alrededores. mediante la digitalización de los registros en una base de datos digitalizada y estructurada. Y haciendo que el proceso de registro de los trabajos realizados sea más rápido y comprensible.

Alcance y contexto

Situación actual: La entidad actualmente realiza los registros diarios de mantenimiento y reparación de forma manual, escrito a mano y almacenados en cuadernos. Dichos registros carecen de una estructura y hasta en algunas ocasiones son complicados de interpretar ya que al no haber un formato específico muchas veces estos registros “omiten” información que otros registros si tienen. Por ejemplo. En algunos registros no se agregan los insumos utilizados y en otros si.



Limitaciones detectadas: Los datos a no tener una estructura y ser escritos manualmente tienen las siguientes limitaciones:

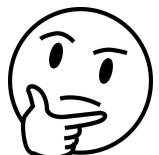
Consumen tiempo: Realizar los registros a mano requieren un tiempo de demora. Muchas veces los trabajadores realizan los registros apuradamente haciendo que sean menos comprensibles.



Los registros son Menos interpretables: Como dijimos anteriormente: Si las personas encargadas de llevar estos registros se encuentran apurados a la hora de registrar los hechos pueden no ser entendibles o puede faltar información ocasionando confusión a cualquier persona que lea estos registros.



Se requiere un mayor esfuerzo al realizar análisis: Debido a que al no ser tan interpretables hace que una persona que quiera realizar un análisis de estos datos. Tenga que esforzarse más.



Propuesta general: Lo que proponemos es una base de datos digital y estructurada cuyos datos ingresan a través de un formulario y son almacenados en una tabla de AirTable. Esto nos permite disminuir el tiempo en el que se demora realizar un registro. hace que los datos almacenados sean más fáciles de interpretar y facilita el desarrollo de futuros análisis de datos.

Alcance y fuera de alcance:

Se incluye		Se excluye	
1	Digitalización de los registros otorgados dándole una estructura	1	Digitalización de otros tipos de registros que no contribuyan al objetivo principal.
2	Limpieza y preprocesamiento con google colab	2	Constantes cambios de estructura de base de datos.
3	Análisis y visualización	3	Realizar otro tipo de base de datos no acordada.
4	Base de datos estructurada	4	Mantenimiento permanente de la base de datos.
5	Mantenimiento y monitoreo inicial.		
6	Indicación de uso del formulario		
7	Capacitación de uso de base de datos.		

Estructura de datos.

La entidad cuenta con una base de datos propia, cuyos datos son de carácter público, en formato papel escritos a mano. Los datos brindados por la entidad son aproximadamente de setecientos registrados manualmente en papel de forma no estructurada. Como los datos son públicos no se requiere ningún tipo de anonimización ya que los registros no comprometen datos personales de las personas que viven en la ciudad de Fernandez y parajes de alrededor.

Metodología aplicada

En esta parte nos dedicamos a nombrar y describir brevemente las etapas realizadas en nuestro proyecto de implementación de base de datos estructural. A continuación se documentan y describen las fases de nuestro proyecto explicando todo el proceso desde la reunión inicial hasta la creación de nuestra base de datos.

1 Análisis del entorno y definición del problema.

En esta etapa se describe la fase inicial del proyecto describiendo brevemente la primera y segunda reunión con nuestro stakeholder para entender el contexto de la entidad actual.

Las reuniones se realizaron con el asesor del área de electricidad. y en la primera reunión se hicieron preguntas para saber:

- Cantidad de empleados.
- cargos y jerarquía
- quienes son los beneficiarios del servicio.
- Si cuentan con bases de datos.

Y obtuvimos información de que hasta el momento el área tenía un equipo de 7 empleados donde 4 de los cuales eran electricistas y 2 eran ayudantes y choferes y uno de ellos era el asesor del área. Y también se descubrió que el servicio prestado no solo abarcaba a la ciudad de Fernandez. sino que también abarcaba a los parajes de alrededor como mistol, la loma. etc. La entidad actual contaba con registros de donde realizaban sus servicios y los insumos utilizados.

La segunda entidad no tenía una problemática en específico pero en la segunda reunión fue para obtener los registros realizados. nos encontramos que los registros se realizaban a mano y detectamos que los registros carecían de una estructura, y que en muchas ocasiones eran escritos con prisa por lo que se volvía difícil entenderlos.

2 Digitalización y limpieza.

En esta etapa los registros se obtienen mediante fotos tomadas por un celular. Los registros capturados y digitalizados abarcan aproximadamente siete meses del servicio realizado que empiezan en diciembre del año 2024 y finalizan en junio del año 2025.

La digitalización se llevó a cabo en una hoja de cálculo de google sheets donde en primera instancia se dividieron las fotos entre cada uno de los integrantes para realizar la digitalización de manera rápida en hojas diferentes y luego los registros digitalizados en una estructura inicial eran fusionados en una sola tabla. La tabla inicial contaba con 14 columnas y 520 registros.

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	fecha	trabajo_realizado	calle	tipo	referencia	barrio	insumos_1	cantidad_1	insumos_2	cantidad_2	insumos_3	cantidad_3	insumos_4	cantidad_4
2	19/02/25	colocacion reflectores	belgrano	publico	ferro	null	lamparas 100w	2	cable concentrico	50mt	reflectores	3	null	null
3	19/02/25	reto de adornos y luminarias	bruno ramirez	publico	pasaje	null	null	null	null	null	null	null	null	null
4	19/02/25	colocacion de lampara	libertad	publico	libertad y guemes	null	lamparas 80w	1	null	null	null	null	null	null
5	19/02/25	colocacion de lampara	libertad	publico	entre guemes y inmigrantes	null	lamparas 100w	1	null	null	null	null	null	null
6	19/02/25	colocacion de lampara	remedios de escalada	publico	entre 9 julio y jesus fernandez	null	lamparas 100w	1	null	null	null	null	null	null
7	19/02/25	colocacion de lampara	moreno	publico	moreno y mayul	independiente	lamparas 100w	1	null	null	null	null	null	null
8	19/02/25	colocacion de lampara	moreno	publico	entre mayul y Gral. taboada	independiente	lamparas 100w	2	null	null	null	null	null	null
9	19/02/25	colocacion	av. del agro	privado	fia gerez	null	corriente	1	null	null	null	null	null	null
10	19/02/25	armado de jaula antibandalos	null	null	taller propio	null	null	null	null	null	null	null	null	null
11	28/01/25	retiro de insumo	san martin	null	municipalidad	null	lamparas 100w	20	null	null	null	null	null	null
12	28/01/25	colocacion de poste y lampara	coronel borges	publico	coronel borges y martin fiero	null	lamparas 80w	1	cable	10mt	poste	1	antibandalos	1
13	28/01/25	colocacion de lampara	sargento cabral	publico	sargento cabral y absalon rojas	null	lamparas 100w	1	null	null	null	null	null	null
14	28/01/25	colocacion de lampara	remedios de escalada	publico	remedios de escalada y 27 de abril	null	lamparas 100w	2	null	null	null	null	null	null

Para la etapa de la limpieza se exporta la nueva tabla digitalizada como archivo csv y se lo carga en un notebook de google colab para la limpieza trabajando la librería pandas. Se realiza un análisis inicial para conocer las columnas y el tipo de columnas se cuentan la cantidad de datos nulos (se escribía null en las columnas donde faltaba o no se entendía ese dato) Y se eliminan columnas cuyos datos nulos superan el 80% de los datos en esas columnas.

para el tratamiento de los datos nulos en las columnas restantes se hace lo siguiente:

- columnas numéricas: los datos nulos son transformados en “0”.
- columnas categóricas o textuales:
 - 1) los valores nulos de las columnas que hacen referencia a lugares se transformaron por “desconocidos”

```
ubicaciones_nulas = ['calle', 'referencia', 'barrio']
df_model[ubicaciones_nulas] = df[ubicaciones_nulas].fillna('desconocido')
```

- 2) los valores nulos que hacen referencia al tipo de insumo utilizado se transformaron a “no se utilizaron insumos”

```
df_model['insumos_1']= df_model['insumos_1'].fillna('no se usaron insumos')
df_model['insumos_1']= df_model['insumos_1'].replace('', 'no se usaron insumos')
df_model['cantidad_1']= df_model['cantidad_1'].fillna(0)
```

- 3) Los valores nulos que hacen referencia a la categoría de trabajo se transformaron en “trabajos generales/ varios”.

```
df_model['categoria'] = df_model['categoria'].fillna("Trabajos generales / Varios")
```

Finalizada la imputación de los datos nulos a continuación se trata el formato de los datos y se unifican aquellos datos que explican una misma información pero distinta manera. Por ejemplo en la columna “calle” la calle “San Martín” estaba escrita en distintas formas como “san martin”, “av san martín” “San martín”. etc.

una vez unificado el formato de las calle lo que se hace es transformar y limpiar las columnas:

- se eliminan los espacios.
- se eliminan las comas.
- se transforman las columnas numéricas en “int”.
- se transforma la columna “fecha” al tipo “datetime”.

El dataset resultante es un dataset limpio de 8 columnas y 512 registros. sin datos nulos y con formato unificado.

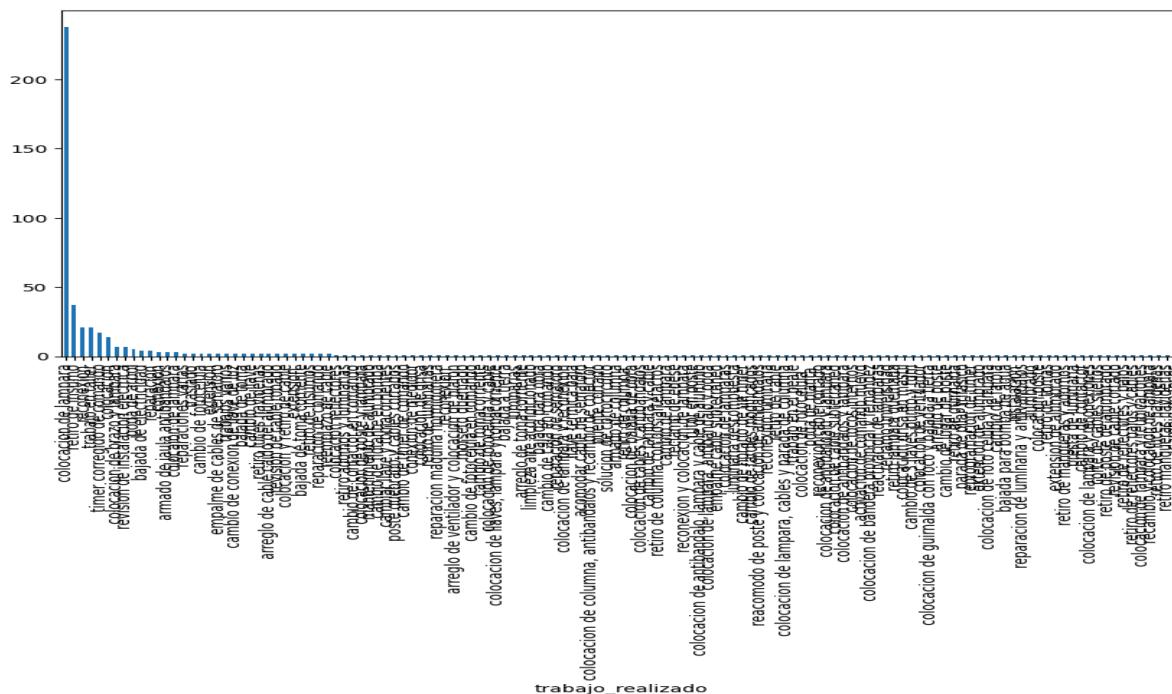
	fecha	trabajo_realizado	calle	categoria	referencia	barrio	insumos_1	cantidad_1
0	19/02/25	colocacion reflectores	belgrano	Instalaciones Varias	ferro	desconocido	lampara 100w	2
1	19/02/25	retiro adornos y luminarias	bruno ramirez	Reparación / Mantenimiento	pasaje	desconocido	no se usaron insumos	0
2	19/02/25	colocacion de lampara	libertad	Instalaciones Varias	libertad y guemes	desconocido	lampara 80w	1
3	19/02/25	colocacion de lampara	libertad	Instalaciones Varias	entre guemes y inmigrantes	desconocido	lampara 100w	1
4	19/02/25	colocacion de lampara	remedios de escalada	Instalaciones Varias	entre 9 julio y jesus fernandez	desconocido	lampara 100w	1
5	19/02/25	colocacion de lampara	moreno	Instalaciones Varias	moreno y mayuli	independiente	lampara 100w	1

3 Análisis exploratorio.

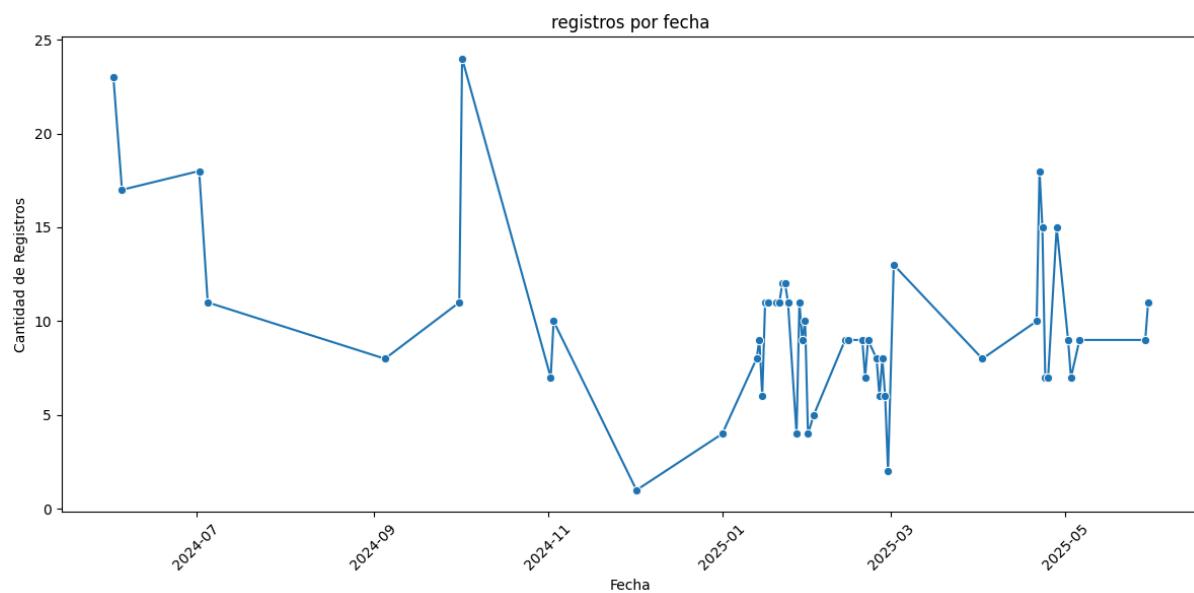
En esta etapa se realizaron gráficos para entender la distribución de datos e intentar descubrir patrones.

En esta etapa nos encontramos que:

Al realizar este gráfico encontramos una gran redundancia en los datos. la cantidad de trabajos realizados eran tan grande que provocaban que el grafico de barras se vuelva inentendible.

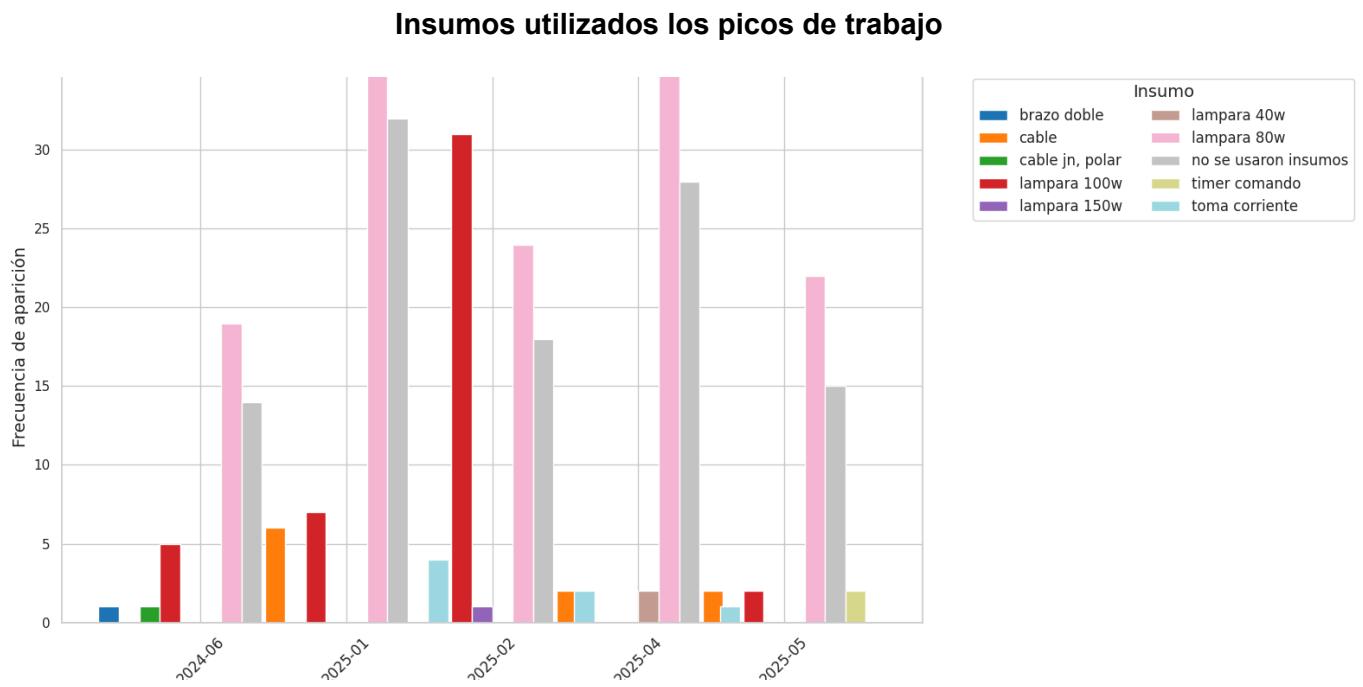


También pudimos encontrar patrones analizando los registros en base a las fechas:



podemos observar que existen picos donde más trabajos se realizaron. por ejemplo entre septiembre y noviembre la actividad tiende a subir a su pico más alto. a partir de noviembre la actividad disminuye hasta diciembre donde vuelve a aumentar la actividad. a partir de enero y en adelante.

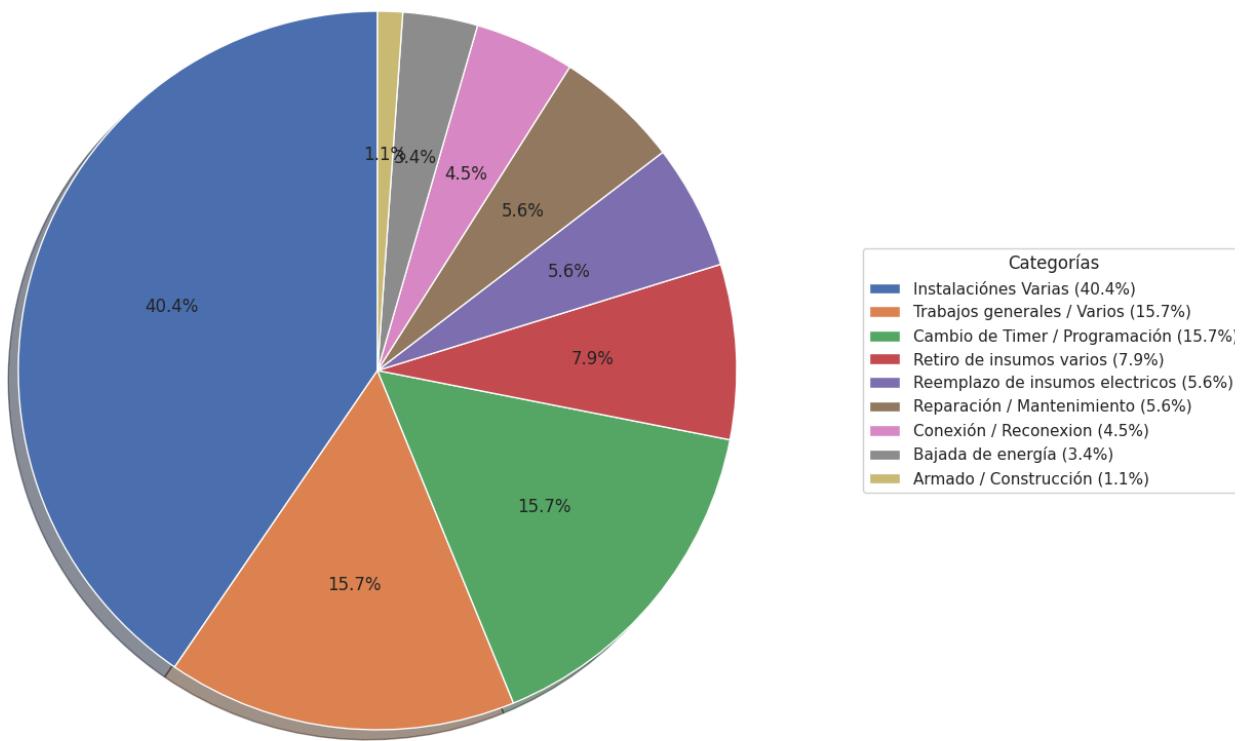
4. Visualizaciones o dashboard



Como primera visualización, hicimos un gráfico de barras múltiple donde el eje x contiene el top 5 meses donde más trabajos se realizó y el eje y que cuenta la frecuencia de insumos. Todo esto con el objetivo de entender cuáles son los insumos que más se consumen en estos cinco meses. Con este gráfico podemos entender que la principal tarea que realiza nuestra entidad es el cambio o colocación de lámparas de 80w del alumbrado público (con una frecuencia mucho mayor a 35). Seguido de colocación de lámparas de 100w para estas fechas y seguido también de trabajos que no necesitan algún insumo.

Trabajos realizados en los años 2024 y 2025

Trabajos realizados en el 2024 (Julio-Diciembre)

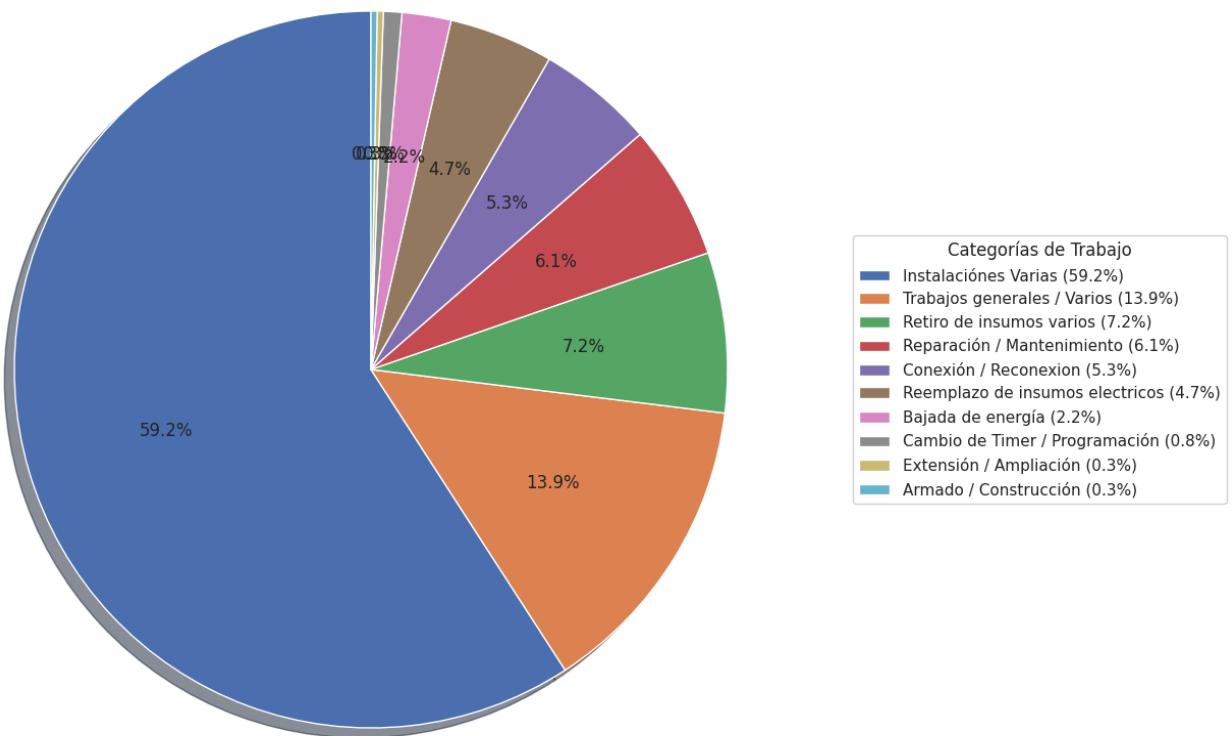


Con el siguiente gráfico de pie podemos observar los tipos de trabajos (categorizados) que se realizaron en el año 2024. Analizando el mes de Julio hasta Diciembre del mismo año.

Se puede observar que:

- un 40% de los trabajos se basaron en hacer instalaciones durante este periodo de tiempo.
- un 15% basado en el cambio/programación de timers.
- un 15% en trabajos generales como podas de árboles, empalme de cables. etc.
- Casi un 8% dedicado al retiro de insumos.
- Casi un 6% en reparación y mantenimiento.
- Un 4,5% en conexiones y reconexiones.
- Un 3,4% en bajadas de energía.
- Y por último un 1% en armado y construcción.

Trabajos realizados en 2025 (Enero-Mayo)



Este gráfico de pie al igual que el otro muestra un porcentaje de los trabajos realizados, pero está basado en el año 2025. Donde se digitalizaron registros de Enero hasta Mayo del mismo año.

Con este gráfico podemos observar que:

- Un 59,2% de los trabajos realizados se basan sobre instalaciones.
- Un 13,9% en trabajos generales.
- Un 7,2% en retiro de insumos.
- Un 6,1% en reparación/mantenimiento.
- Un 5,3% en conexiones y reconexiones.
- Un 4,7% en reemplazo de insumos.
- Un 2,2% en bajada de energía.
- Un 0,8% en cambios de timer
- Un 0,3% basado en extensión, y armado.

Solución desarrollada

Descripción general de la solución técnica.

La solución se enfoca en la digitalización y estructuración del registro histórico de refacciones del servicio eléctrico de la Municipalidad de Fernández, que se manejaba en cuadernos de papel. El proyecto incluyó el análisis exploratorio de estos datos para optimizar la planificación de tareas y la asignación de recursos. Adicionalmente, se propuso e implementó un nuevo sistema de registro digital basado en Airtable para la gestión de futuros trabajos. Este nuevo flujo de trabajo permite al cliente acceder y editar la base de datos de forma simultánea, y facilita la búsqueda en el registro histórico.

Herramientas y librerías utilizadas

Recolección/Tabulación Inicial: Hojas de cálculo de **Google Sheets** para la digitalización inicial de los registros en papel, creando la estructura inicial.

Limpieza y Procesamiento de Datos:

Python y la librería **Pandas** dentro de un *notebook* de **Google Colab** para la limpieza, normalización, imputación de nulos y preprocesamiento de los datos históricos.

Análisis y Visualización:

Python (con librerías como Matplotlib) para realizar el análisis exploratorio de datos (EDA) y crear gráficos de distribución temporal y frecuencia de trabajos/insumos.

Software para el Nuevo Registro Digital:

Airtable fue seleccionado para la nueva **Base de Datos Mayor** estructurada, que se alimenta mediante un formulario de registro.

Componente de ML Propuesto: El **Chatbot** conceptual representa un componente futuro que usaría **Procesamiento de Lenguaje Natural (PLN)** para capturar reclamos.

Flujo de trabajo:

El flujo se divide en el Procesamiento Histórico y la Nueva Dinámica de Producción (Flujo de Producción):

Procesamiento de Datos Históricos (Análisis del proyecto):

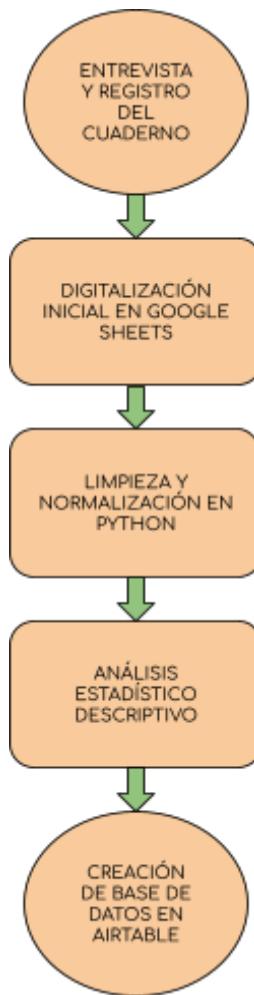
Entrevista inicial y captura de fotos del cuaderno

Digitalización inicial en **Google Sheets**

Limpieza y normalización en **Google Colab (Python/Pandas)**

Análisis Estadístico Descriptivo

Consolidación en la **Base de Datos Mayor (Airtable)**.



Nueva Dinámica de Registro (Flujo de Producción Propuesto):

Captura del Reclamo (ML/AI): Los usuarios hacen reclamos a través del **Chatbot**. El Chatbot interpreta (ML/PLN) la solicitud y la registra en la **Base de Datos Menor (Reclamos Pendientes)**.

Ejecución del Trabajo: Los empleados toman el reclamo de la Base de Datos Menor, realizan la obra.

Registro Final: El empleado completa un **formulario** ([enlace de Airtable](#)) con el detalle de la obra realizada.

Consolidación: El formulario alimenta directamente la **Base de Datos Mayor** estructurada y digitalizada en **Airtable**, completando el ciclo del registro histórico.

Análisis estadístico realizado.

El principal entregable analítico del proyecto fue el Análisis Estadístico Descriptivo (EDA) y Exploratorio de Datos de los registros históricos. Los principales hallazgos fueron:

Ánálisis de Frecuencia de Trabajos: Identificación de la "**colocación de lámpara**" como la actividad predominante.

Ánálisis de Insumos en Picos: Se visualizó que el principal insumo consumido en períodos de alta demanda es la **lámpara de 80w**.

Categorización de Trabajos: Se agruparon los trabajos en categorías (Instalaciones Varias, Trabajos generales, etc.), observando que las "**Instalaciones Varias**" representan la mayor parte de los trabajos.

Componente de Machine Learning (ML) proyecto futuro:

La implementación de un **Chatbot** para la recepción de reclamos se propone como la aplicación de Machine Learning (mediante **Procesamiento de Lenguaje Natural - PLN**). El chatbot no solo captura el reclamo en texto libre, sino que lo interpreta, clasifica (ej: "falla de alumbrado", "cable caído"), y lo estructura automáticamente para alimentar la **Base de Datos Menor**. Esto optimiza la entrada de datos, reduce errores y agiliza la gestión del servicio.

8. Propuesta de base de datos

Entidad (Tabla)	Atributos Clave (Columnas)
Trabajos Realizados (Principal)	Fecha, Trabajo Realizado, Colle [FK], Categoría, Referencia, Barrio [FK], Insumo Utilizado [FK]
Calles	Nombre de la Calle (PK), Barrio [FK], Cantidad Total de Trabajos
Barrios	Nombre del Barrio (PK), Cantidad de Trabajos, Última Fecha de Trabajo
Insumos	Nombre del Insumo (PK), Descripción Resumida, Cantidad en Stock, Total Consumido

Relaciones (Cardinalidad)

Trabajos Realizados ↔ Calles

Relación: N:1 (Muchos Trabajos a Una Calle)

Un mismo trabajo se realiza en una sola calle. Sin embargo, una Calle puede tener asociados muchos registros de Trabajos Realizados.

La tabla Trabajos Realizados contiene la clave foránea (Calle).

Trabajos Realizados ↔ Barrios

Relación: N:1 (Muchos Trabajos a Un Barrio)

Un mismo trabajo se realiza en un solo Barrio. No obstante, un Barrio puede tener asociados muchos registros de Trabajos Realizados.

La tabla Trabajos Realizados contiene la clave foránea (Barrio).

Trabajos Realizados ↔ Insumos

Relación: N:1 (Muchos Trabajos a Un Insumo)

Un trabajo registrado utiliza un tipo de Insumo (el campo en la tabla principal es Insumo Utilizado). Un único Insumo (ej: lámpara de 80w) puede ser el Insumo Utilizado en muchos registros de trabajos.

La tabla Trabajos Realizados contiene la clave foránea (Insumo Utilizado).

Calles ↔ Barrios

Relación: N:1 (Muchas Calles a Un Barrio)

Justificación: Una Calle pertenece a un único Barrio , pero un Barrio contiene muchas Calles.

Mapeo: La tabla Calles contiene la clave foránea (Barrio).

Justificación del Diseño Lógico (Normalización)

El diseño de la base de datos implementado en Airtable sigue los principios de la Tercera Forma Normal (3NF):

Eliminación de Redundancia y Mantenimiento de la Coherencia (Normalización)

Creación de Tablas de Referencia (Dimensiones): En lugar de repetir datos en cada registro de Trabajos Realizados (la tabla transaccional) , se crearon tablas de catálogo separadas (Barrios, Calles, Insumos).

Si la Calle "San Martín" cambia de nombre, solo se debe actualizar un registro en la tabla Calles , en lugar de cientos de registros en la tabla Trabajos Realizados. Esto garantiza la coherencia de los datos y minimiza los errores de tipografía (como se encontró en el análisis histórico: "san martin", "av san martín", "San martín").

Implementación y enlace al archivo.

AIRTABLE

CAPTURAS DE AIRTABLE:

Trabajos Realizados

Fecha	Trabajo Realizado	Calle	Categoría	Referencia	Barrio
2025-02-28	Abrir > colocacion de lampara	rafael ruiz	Instalaciones Varias	rafael ruiz y belgrano	desconocido
2025-01-20	colocacion de lampara	uriarte	Instalaciones Varias	uriarte y matienzo	desconocido
2025-07-02	colocacion de lampara	independencia	Instalaciones Varias	independencia y 20 junio	desconocido
2024-12-31	colocacion de lampara	garcia mendez	Instalaciones Varias	garcia mendez y belgrano	desconocido
2025-03-02	cambio de focos	av. 25 de mayo	Reemplazo de insumos el...	veredon y bulevar	desconocido

Calles

Nombre de la Calle	Barrio	Cantidad Total de Trabajos	Última Fecha de Trabajo
cordoba	la loma	9	0
lugones	2 de abril	1	0
av.del agro	quemado	1	0
25 de mayo	cooperativa	1	0
av san martin	21 de septiembre	2	0

Barrios

Nombre del Barrio	Foto del Barrio	Cantidad de Trabajos...	Última Fecha de Trabajo
desconocido		367	0
30 viviendas		3	0
hospital		3	0
la loma		4	0
villa elisa		3	0

Insumos

Nombre del Insumo	Descripción Resumida (AI)	Cantidad en Stock	Total Consumido
- brazo doble	Los campos...	0	6
Unidad toma corriente	Puntos de conexión...	100	10
- postes	Los campos...	0	2
- lampara 40w	Los campos...	0	12
- antibandalo	Los campos...	0	1
- alumbrado	Los campos...	0	150
- farola 40w	Los campos...	0	10
- lamparas 150w	Los campos...	0	4

9.Documentación en GitHub

La documentación del proyecto se estructuró siguiendo buenas prácticas recomendadas para trabajos de ciencia de datos, análisis exploratorio y desarrollo colaborativo. El repositorio de GitHub funciona como un punto central donde se almacena tanto el código como los datos, permitiendo reproducibilidad, claridad y trazabilidad del proceso.

El objetivo de esta organización es garantizar que cualquier persona que acceda al repositorio (docente, compañero o revisor externo) pueda entender el flujo completo del proyecto desde la obtención del dataset inicial hasta la generación del informe final. Para ello, se adoptó una arquitectura basada en directorios bien definidos, cada uno con un propósito concreto.

Para esta instancia se decidió usar una estructura que cuenta con los siguientes archivos y carpetas:

1. [README.MD](#) (Procurando nombrar el proyecto, dar el contexto inicial, objetivo y un resumen de lo realizado en los distintos entornos y plataformas utilizadas).
2. Directorio/Notebooks: donde se encontrará el Notebook realizado en Google Colab en un formato .ipynb
3. Directorio/Data/Raw: Donde se colocará el primer dataset realizado justo después de la digitalización en formato .CSV (el dataset sin procesar).
4. Directorio/Data/Processed: En esta carpeta se colocará el dataset final en formato .CSV (luego de la limpieza y procesamiento).
5. Directorio/Docs: Esta carpeta está destinada a contener este mismo documento en formato .pdf
6. Directorio/Assets: Capturas de la base de datos en Airtable y de algunos gráficos del dashboard.
7. Requirements.txt: En este archivo deberemos colocar las librerías exactas utilizadas en Google Colab a la hora de realizar la limpieza (Ej. numpy, pandas ==2.1.4, etc.)

10. Conclusiones y Próximos Pasos

Lo que logramos concretar en esta etapa fue un avance fundamental para la Municipalidad de Fernández. Pudimos rescatar y sanear el registro histórico de refacciones eléctricas, que antes estaba disperso en cuadernos escritos a mano, y lo transformamos en una base de datos digital ordenada y confiable. Este proceso incluyó la digitalización de cerca de setecientos registros y un trabajo exhaustivo de normalización de calles y fechas realizado en Python y Google Colab.

Además, uno de los logros centrales fue el Análisis Estadístico Descriptivo, que nos permitió identificar patrones claros, como el aumento de trabajos entre septiembre y noviembre y la alta frecuencia de colocación de lámparas de 80 W. También dejamos implementada una solución moderna para el registro de nuevos casos en Airtable, lo que permite un ingreso estructurado mediante formularios, reduciendo errores y mejorando la organización general del área.

En cuanto a las mejoras futuras, la prioridad sería la implementación de la Base de Datos Menor conectada a un chatbot, capaz de interpretar reclamos en lenguaje natural mediante técnicas de Machine Learning, clasificar el tipo de falla y registrar automáticamente la ubicación. Esto permitiría reducir tiempos de respuesta y minimizar errores humanos.

Descubrimos que un proyecto de datos empieza mucho antes de escribir el primer script: de hecho, lo más desafiante fue la etapa de acercamiento, entrevistas y digitalización manual de registros caóticos. Esto nos enseñó el valor de la normalización, especialmente al transformar descripciones ambiguas como “al lado de la casa de Patricia” en datos útiles y georreferenciables.

Finalmente, entendimos cómo un análisis estadístico básico pero bien hecho puede generar un impacto directo y positivo en la gestión de recursos públicos, haciendo el trabajo más eficiente, transparente y efectivo.