



IMPLEMENTACION MACHINE LEARNING

Para análisis de datos de Medicion de Gas



¿Conocían a
MERLIN?



1

2

3



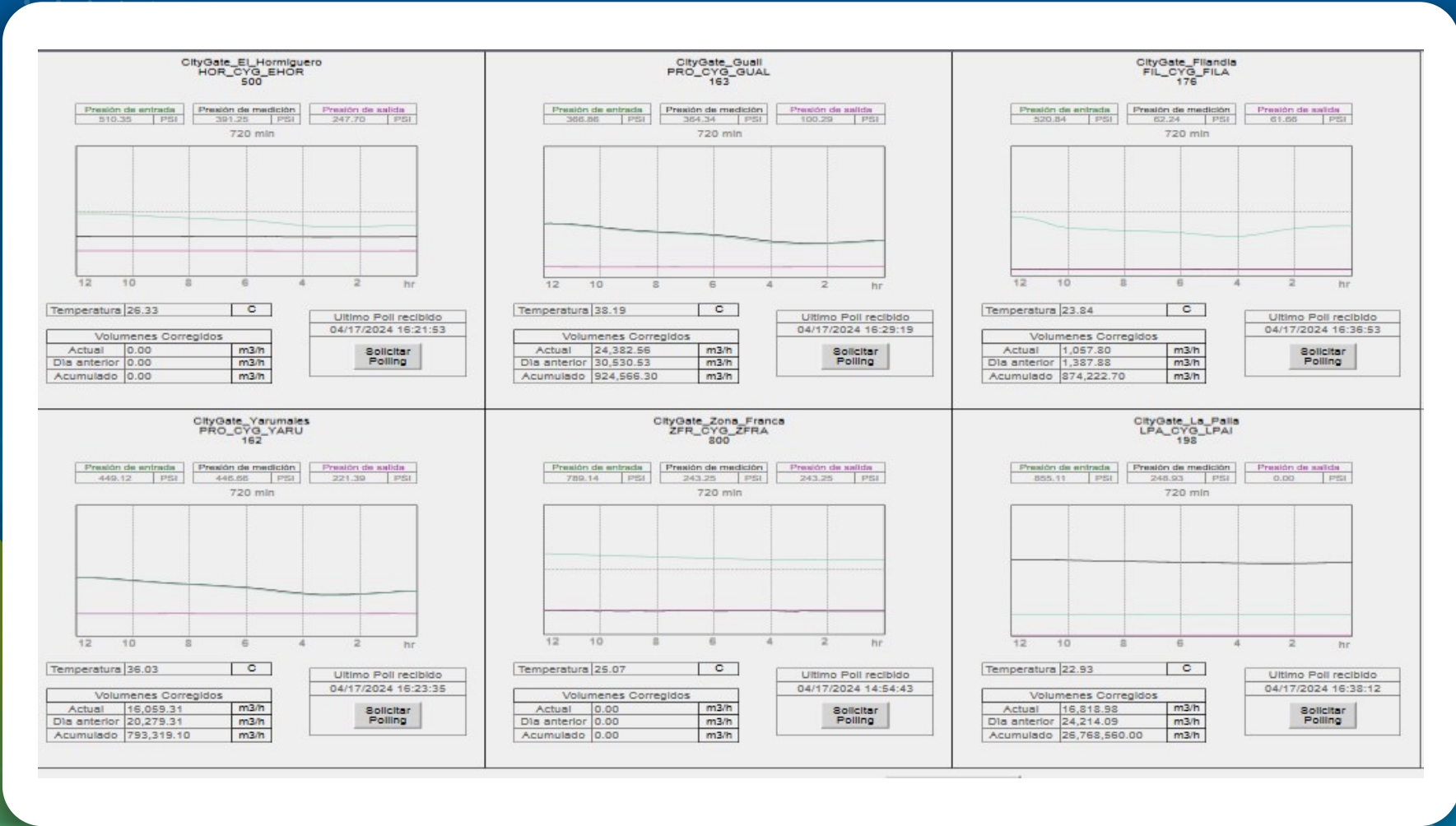
Nuestra plataforma de trabajo



Cantidad y Calidad de datos



Herramienta de Eficiencia





```
# prompt: crear df a partir de merged_df que me calcule el promedio, percentil 25, percentil 75, rango intercuartil de Vol, agrupando por id_sitio, dia_semana y Hora_Modelo

promedio_df_Por = merged_df_sin_NaN.groupby(['id_sitio', 'dia_semana', 'Hora_Modelo'])['%Hora'].mean().reset_index(name='Promedio_Por')
percentil_25_df_Por = merged_df_sin_NaN.groupby(['id_sitio', 'dia_semana', 'Hora_Modelo'])['%Hora'].quantile(0.25).reset_index(name='Percentil_25_Por')
percentil_75_df_Por = merged_df_sin_NaN.groupby(['id_sitio', 'dia_semana', 'Hora_Modelo'])['%Hora'].quantile(0.75).reset_index(name='Percentil_75_Por')
#rango_intercuartil_df = percentil_75_df.merge(percentil_25_df, on=['id_sitio', 'dia_semana', 'Hora_Modelo'])['Percentil_75'] - percentil_25_df['Percentil_25']
#rango_intercuartil_df = rango_intercuartil_df.rename(columns={0: 'Rango_Intercuartil'})

df_final_Por = promedio_df_Por.merge(percentil_25_df_Por, on=['id_sitio', 'dia_semana', 'Hora_Modelo'])
df_final_Por = df_final_Por.merge(percentil_75_df_Por, on=['id_sitio', 'dia_semana', 'Hora_Modelo'])
df_final_Por["rango_intercuartil_Por"] = df_final_Por["Percentil_75_Por"] - df_final_Por["Percentil_25_Por"]
df_final_Por["Limite_Superior_Leve_Por"] = df_final_Por["Percentil_75_Por"] + 1.5*df_final_Por["rango_intercuartil_Por"]
df_final_Por["Limite_Superior_Extremo_Por"] = df_final_Por["Percentil_75_Por"] + 3*df_final_Por["rango_intercuartil_Por"]
df_final_Por["Limite_Inferior_Leve_Por"] = df_final_Por["Percentil_25_Por"] - 1.5*df_final_Por["rango_intercuartil_Por"]
df_final_Por["Limite_Inferior_Extremo_Por"] = df_final_Por["Percentil_25_Por"] - 3*df_final_Por["rango_intercuartil_Por"]
df_final_Por["Limite_Inferior_Leve_Por"] = df_final_Por["Limite_Inferior_Leve_Por"].apply(lambda x: 0 if x < 0 else x)
df_final_Por["Limite_Inferior_Extremo_Por"] = df_final_Por["Limite_Inferior_Extremo_Por"].apply(lambda x: 0 if x < 0 else x)

from IPython.display import display

display(df_final_Por)
#print(df_final.head())
```

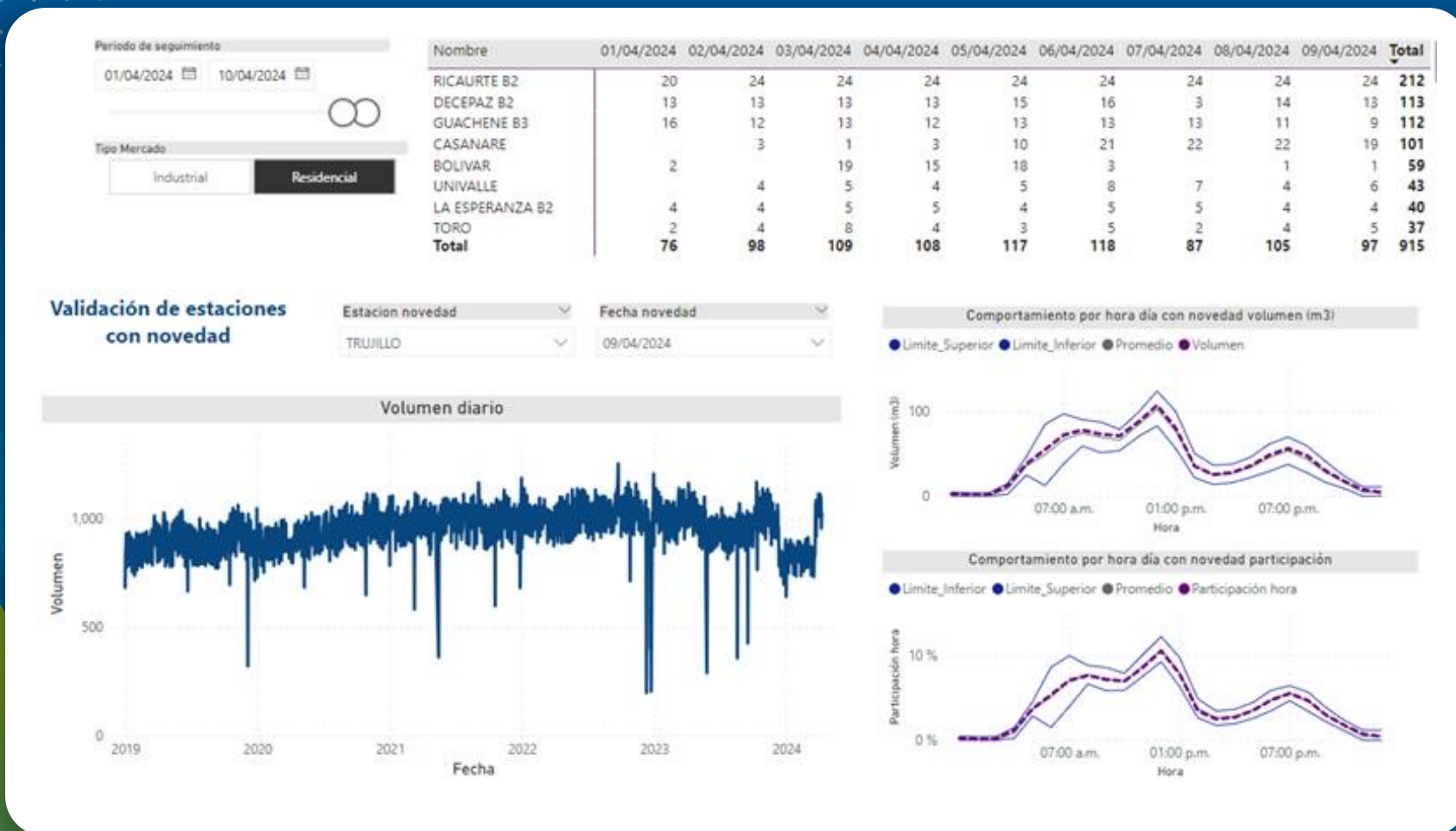
+ 3M datos

+ 290

estaciones
de GN y GNC

MERLIN

NUESTRO RESULTADO

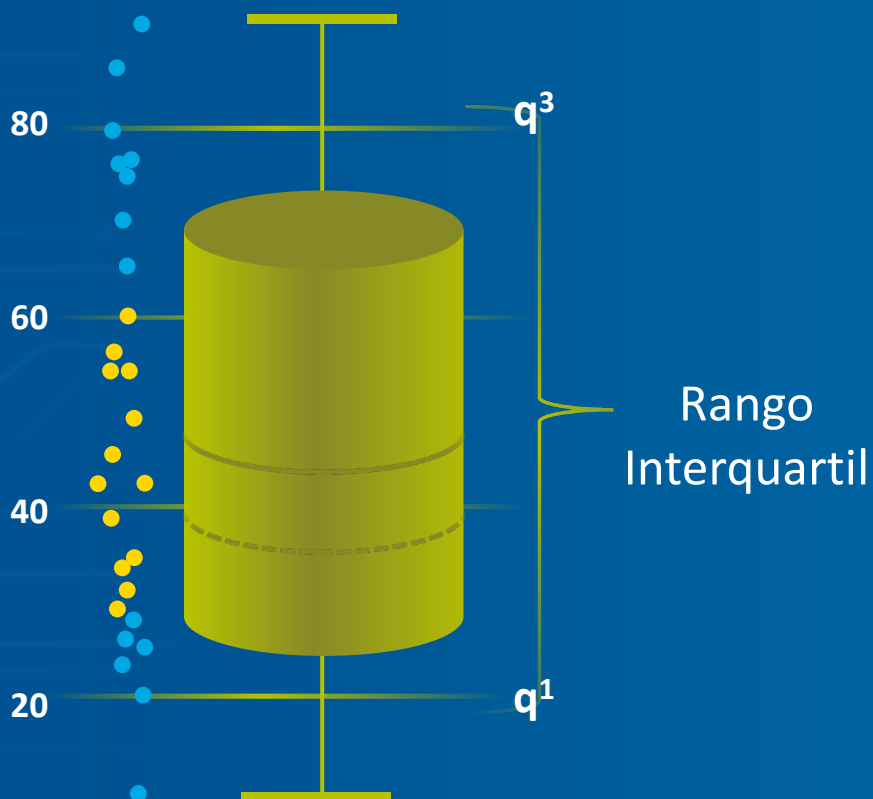


BASE DE DATOS

El 25% de los datos están por encima de q_3 .

En la caja encontramos el 50% de nuestros datos.

25% de los datos están por debajo de q_1 .



Compilación de datos históricos sobre variables de medición.

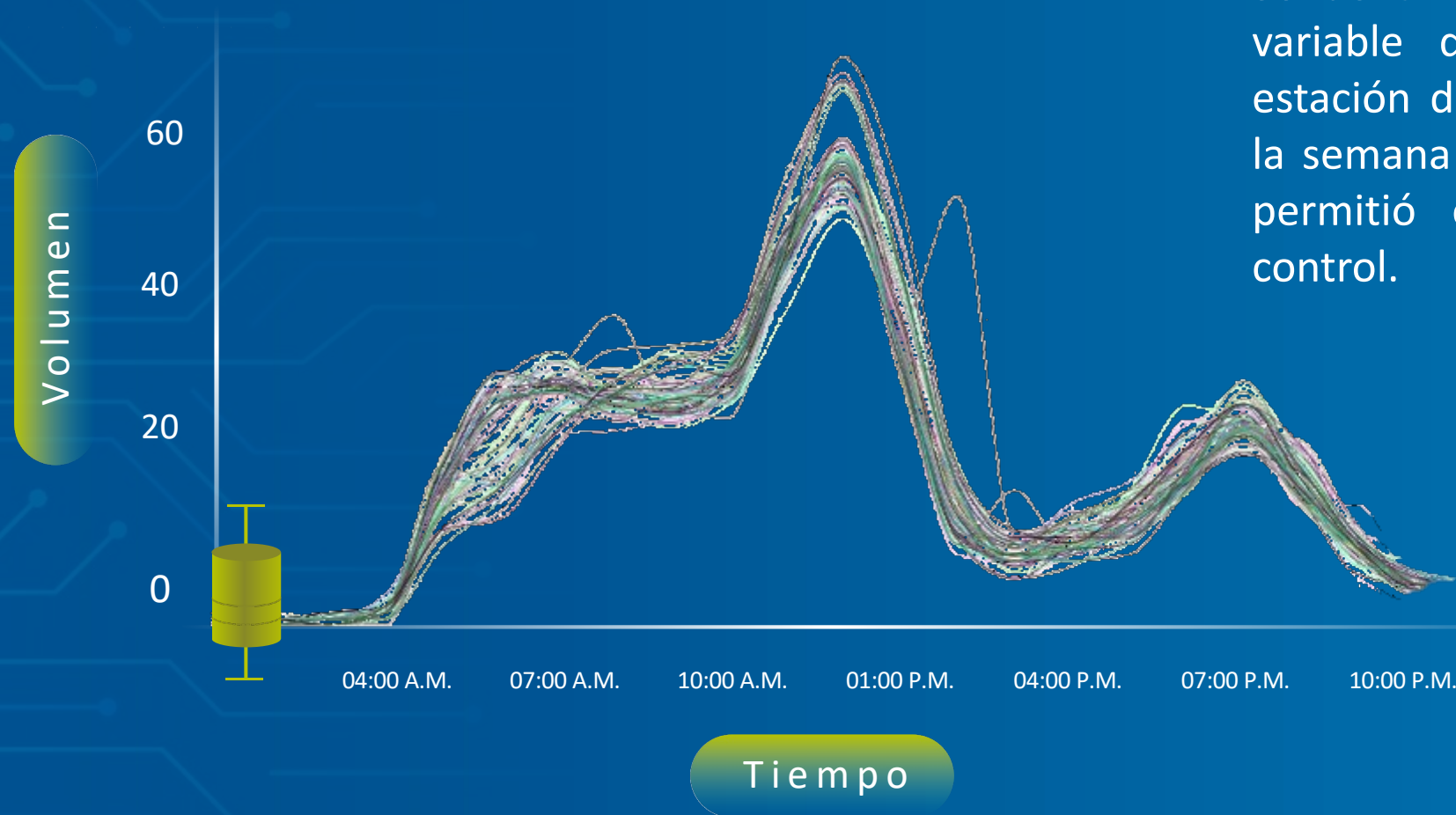


Establecer límites objetivo para tomar acciones.



Identificar anomalías que afectan el balance de gas.

COMO FUNCIONA?

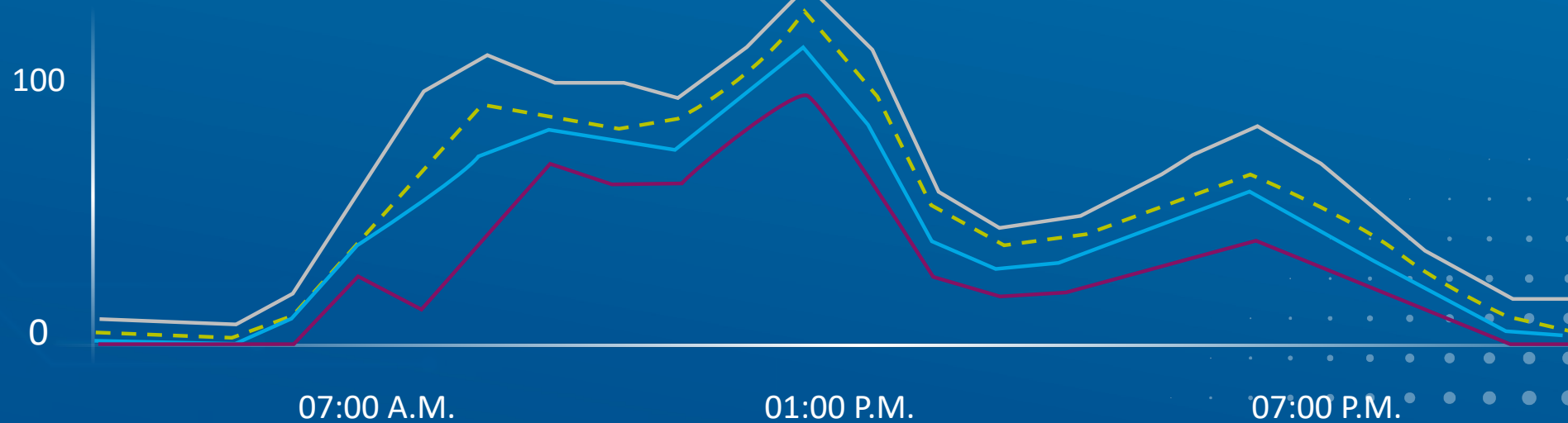


Se identifica La distribución de la variable de volumen en cada estación discriminando el día de la semana y hora del día, lo que permitió establecer límites de control.

Comportamiento del volumen por hora

- Limite superior
- Limite inferior
- Promedio
- Volumen real

Volumen

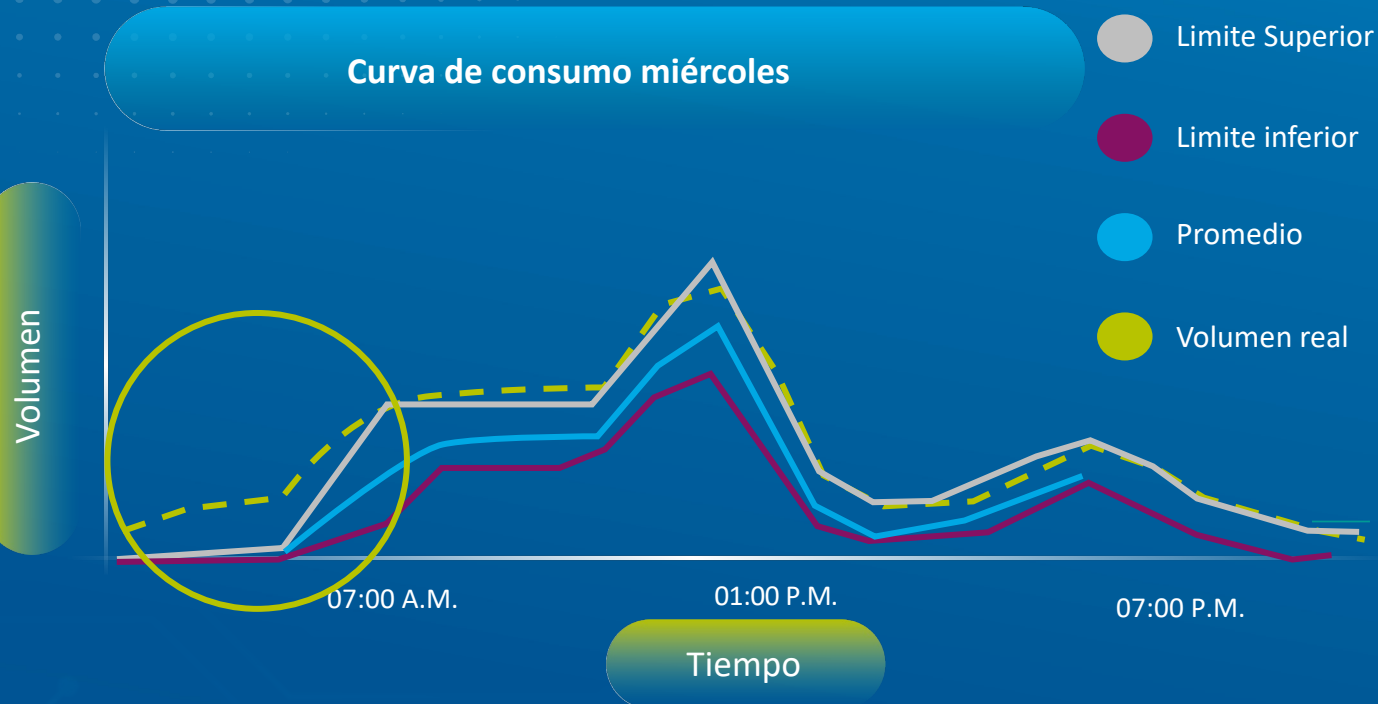


Tiempo

CASO DE ÉXITO I

Localidad: Bolívar
Usuarios: 1.527

Curva de consumo miércoles



Acciones tomadas



Verificación de nuevos usuarios y consumo comercial.



Patrullaje de redes de distribución



Validación de todo el Sistema de medición



Validación de emergencias presentadas

Identificamos un nuevo usuario comercial que inicio producción en la zona con comportamiento atípico.

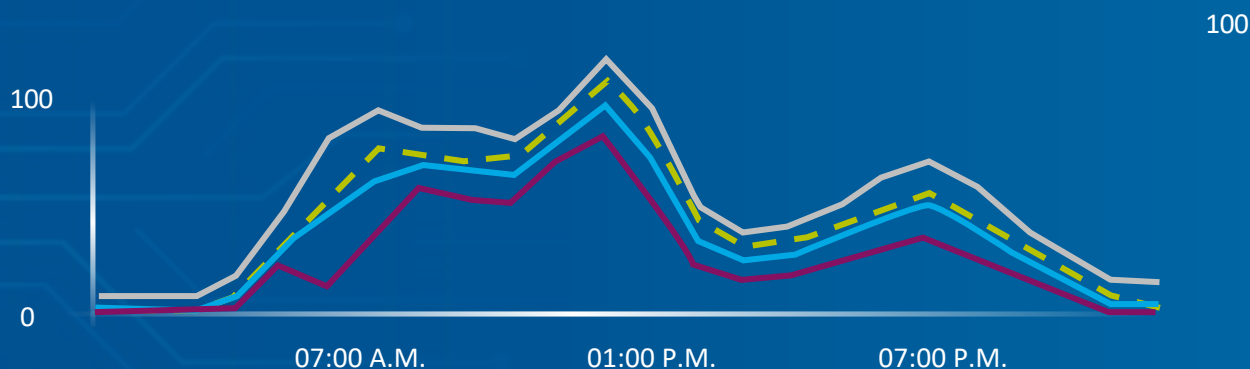
CASO DE ÉXITO II

Localidad: Trujillo
Usuarios: 2.213



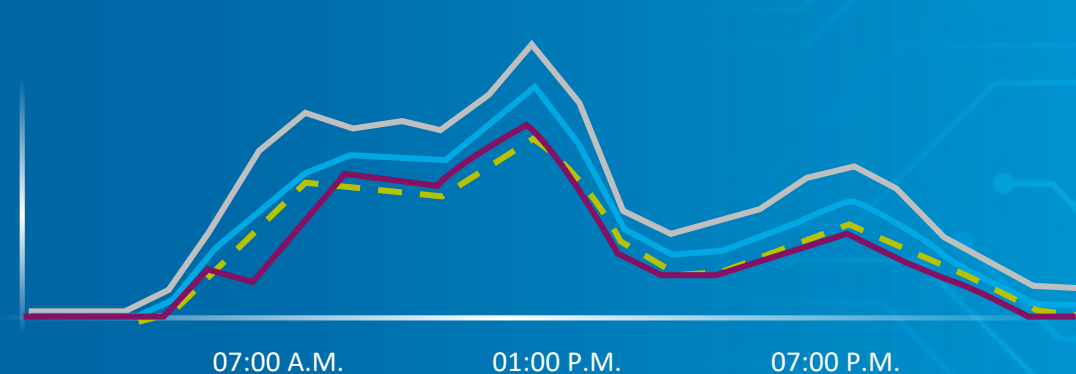
Curva de consumo un día típico

Volumen



Tiempo

Curva de consumo un día atípico



Tiempo

Encontramos una falla en el sensor de temperatura que estaba afectando la corrección del volumen.

BENEFICIOS



Monitoreo del gas
medido



Identificar gas no
medido



Seguridad en la
operación



Mantener el nivel de
perdidas por debajo
del 1%



Precisión en la
medición



**Esta iniciativa forma parte de la estrategia
carbono neutro 2040**

NUESTRA VISION



1

Integrar otras variables operativas monitoreadas desde el sistema **SCADA** y replicarlas para otras áreas de la empresa

2

Integrar Merlin con el sistema SCADA para el monitoreo en **tiempo real** de las variaciones, proporcionando asistencia al equipo de centro de control

3

Incorporar inteligencia artificial para mejorar las **aplicaciones** y detección de anomalías

Participación AGA 2024





SOMOS ENERGIA QUE IMPULSA EL DESARROLLO

