

**Proyecto reporte entregable #1 – Caso Electro Dunas**

• *Problema que abordarán y su contexto.*

Electro Dunas, es una compañía de distribución de electricidad que opera en el Perú, en los departamentos de Ica, Huancavelica y Ayacucho. La compañía atiende a clientes en dos grandes segmentos: por un lado, el mercado regulado, donde hay una tarifa estándar para todos los abonados y otro grupo de clientes no regulados, quienes tienen tarifas negociadas y son usuarios industriales con consumos importantes, la empresa los define como consumos “mayores a 2.5 MW/año/punto de suministro”.

En el segmento no regulado, la empresa ha tenido un crecimiento considerable de sus ventas, +43.8% de crecimiento año a año entre 2021 y 2022. Ante esta circunstancia Electro Dunas requiere tener más control sobre los costos de la operación en la distribución a clientes no regulados, en especial, sobre las pérdidas no técnicas de electricidad. En el negocio de transmisión de energía hay dos tipos de pérdidas en las redes, técnicas y no técnicas. Las primeras están relacionadas con las ineficiencias que existen en los sistemas de transmisión de electricidad (pérdidas en el cableado, transformadores y otros componentes de la red que pueden perder parte de la electricidad que transmiten por restricciones técnicas). Mientras que las pérdidas no técnicas están relacionadas con comportamientos fraudulentos de los clientes, en los que hay manipulación de los equipos para tener una menor facturación.

• *Pregunta de negocio y alcance del proyecto.*

Electro Dunas requiere identificar qué clientes no regulados presentan pérdidas no técnicas de electricidad en su negocio de distribución, para lo cual requiere una solución basada en analítica que le permita visualizar los datos relacionados con el consumo de cada uno de sus clientes, así como identificar consumos anómalos que puedan ser indicativos de comportamientos fraudulentos, de modo que pueda reducir la mayor cantidad posible de costos asumidos por la empresa por factores no técnicos.

A partir de un exitoso ejercicio de detección de anomalías en los consumos de los clientes no regulados, se espera identificar y clasificar adecuadamente a los clientes entre potenciales generadores de pérdidas no técnicas y clientes con consumos normales, orientando a Electro Dunas sobre cómo gestionar a sus clientes.

Este proyecto corresponde al universo de problemas de aprendizaje no supervisado, por lo que el éxito del proyecto se podría medir posteriormente por Electro Dunas con un indicador de qué porcentaje (%) de pérdidas no técnicas de electricidad se logra identificar a través de la herramienta.

- *Descripción de conjuntos de datos a emplear.*

El conjunto de datos principal corresponde a un panel para 30 clientes no regulados de Electro Dunas, con el registro de consumo energía por hora, entre el 01 de enero de 2021 y el 31 de marzo de 2023, por lo que para cada cliente se tienen 19.681 observaciones para la energía activa (kWh), energía reactiva (kVarh) y voltaje:

```
datos > DATOSCLIENTE3.csv
1 Fecha,Active_energy,Reactive_energy,Voltaje_FA,Voltaje_FC
2 2021-01-01 00:00:00,1.1402510106302337,0.8036245713000879,894.3714487006351,850.7576394346074
```

**Energía activa entregada:** representa la cantidad de energía eléctrica real entregada en un período de tiempo específico, generalmente se mide en kilovatios-hora (kWh). En otras palabras, es la energía que realiza trabajo efectivo, como encender luces, hacer funcionar motores o calentar dispositivos. Es la que se usa para el cálculo del cobro en para cada periodo de facturación.

**Energía reactiva entregada:** indica la cantidad de energía intercambiada entre la fuente de energía y una carga sin realizar trabajo útil. Esta energía se debe principalmente a la capacitancia e inductancia de los equipos eléctricos y no produce trabajo real, pero es esencial para mantener la calidad del voltaje en el sistema eléctrico. Se mide en kilovatios-ampere reactivos-hora (kVARh).

Esta energía se transporta a través de las líneas de alta tensión en corriente alterna, por lo que no toda la electricidad que consumimos es energía activa. Al contrario, la energía reactiva le quita parte de espacio en los sistemas de transporte, pues hay que cubrir la demanda, aunque no tenga ninguna utilidad.

**Voltaje o Tensión:** Diferencia de potencial (también denominada voltaje) es una magnitud física que cuantifica la diferencia de potencial eléctrico entre dos puntos.

Adicionalmente, se cuenta con una tabla auxiliar de datos complementarios que describen la actividad de cada cliente de acuerdo al sector económico al que pertenece.

- *Exploración de los datos.*

A partir de una muestra de los datos, se construyó la exploración inicial de los datos:

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 98405 entries, 0 to 98404
Data columns (total 5 columns):
 #   Column          Non-Null Count  Dtype  
---  -
 0   Fecha           98405 non-null  object  
 1   Active_energy    98405 non-null  float64  
 2   Reactive_energy  98405 non-null  float64  
 3   Voltaje_FA       98405 non-null  float64  
 4   Voltaje_FC       98405 non-null  float64  
dtypes: float64(4), object(1)
memory usage: 3.8+ MB
```

La base de datos de consumo para 5 clientes tiene 5 columnas y 98405 observaciones. Las columnas son: la fecha de observación del consumo, y 4 variables numéricas correspondientes a las 4 métricas

de consumo que se describieron anteriormente. En ninguna observación hay información nula. A continuación se hace una descripción de la distribución del registro de consumo cada hora para cada una de las 4 métricas, de los clientes no regularizados.

```
[25] data.describe()
```

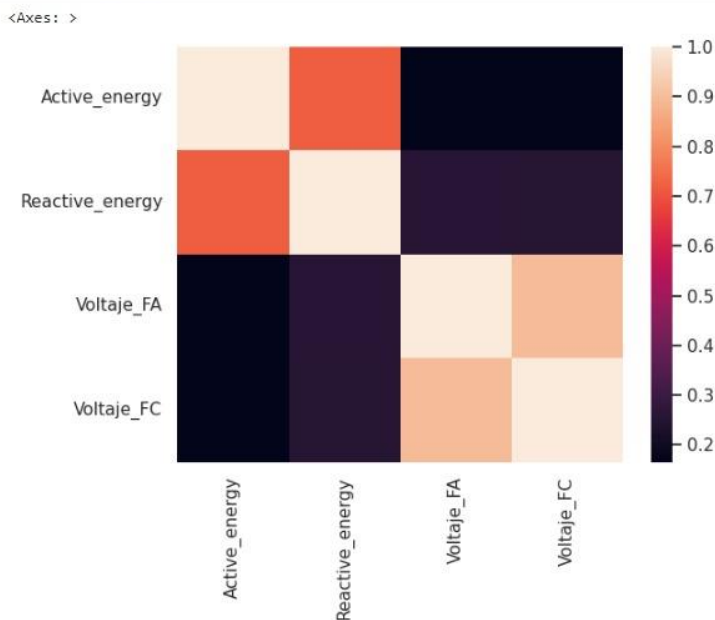
	Active_energy	Reactive_energy	Voltaje_FA	Voltaje_FC
count	98405.000000	98405.000000	98405.000000	98405.000000
mean	2.659339	1.687157	615.392029	615.483875
std	2.099133	1.473453	202.242255	203.508804
min	0.000240	0.000051	146.813497	184.697528
25%	1.176934	0.655413	483.064227	476.315429
50%	2.056246	1.399884	600.594789	603.643077
75%	3.374902	2.083070	699.786733	699.680374
max	14.622644	11.135141	1704.454090	1740.995891

Por último, se hizo un breve análisis de la correlación en el comportamiento de las 4 métricas de consumo de energía, donde se destacan la fuerte relación positiva que existe entre las dos variables de voltaje y la relación entre las variables energía activa y reactiva:

```
[38] import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt

sns.set_theme(style="white")
corr = data.corr()

sns.heatmap(corr)
```



- *Maqueta del prototipo.*

A continuación se muestra el esquema de un tablero de control, el cual permite la visualización de los datos correspondientes al consumo de energía para clientes no regularizados y la solución analítica para la detección de pérdidas no técnicas.



El tablero lleva en el encabezado un título el cual permite identificar el problema a monitorear. En la parte superior derecha se encuentra el listado de los ID de los clientes que según el modelo de detección de anomalías, se presume, efectúan pérdidas no técnicas en el consumo de energía. En la parte superior izquierda del tablero se encuentran 4 botones que permiten filtrar la información por medición (energía reactiva, energía no reactiva, voltaje), sector económico, fecha de observación y cliente sobre el cual se mostrara en la parte derecha del tablero una gráfica del comportamiento por horas de la métrica filtrada del cliente seleccionado en una ventana de observación según las fechas elegidas, junto con un tablero de control que muestra el comportamiento más reciente versus 4 periodos de tiempo. Adicionalmente en la parte inferior derecha del tablero se graficarán los comportamientos históricos, promedios de las 4 métricas para los clientes del sector elegido, esto con el fin de poder hacer una comparación. Finalmente, en la parte inferior izquierda, se construirá un tablero que permita monitorear el comportamiento del modelo de detección de anomalías.

- *Repositorios con todo el código (capturas y enlaces que soporten la creación y uso de repositorios)*

Se crea el repositorio **Electro Dunas** en GitHub, como se presenta a continuación:



Imagen que muestra el proyecto y la estructura base.

URL: <https://github.com/juanrave/electrodunas>

- *Conjuntos de datos a emplear.*

El conjunto de datos a emplear corresponde a una muestra de 30 clientes no regulados, información que es generada en un archivo por cada cliente, como se muestra a continuación.

- *Captura de DVC de los archivos.*

Para el control de las versiones de los datos, se instaló DVC en una maquina con sistemas operativo Ubuntu sobre AWS.

```
(env-dvc) ubuntu@ip-172-31-24-123:~/electroDunas$ dvc init
Initialized DVC repository.
```

You can now commit the changes to git.

DVC has enabled anonymous aggregate usage analytics.  
Read the analytics documentation (and how to opt-out) here:  
<<https://dvc.org/doc/user-guide/analytics>>

What's next?

- Check out the documentation: <<https://dvc.org/doc>>
  - Get help and share ideas: <<https://dvc.org/chat>>
  - Star us on GitHub: <<https://github.com/iterative/dvc>>
- ```
(env-dvc) ubuntu@ip-172-31-24-123:~/electroDunas$
```

Imagen que presenta la inicialización del repositorio DVC en el ambiente del servidor Ubuntu.

```
(env-dvc) ubuntu@ip-172-31-24-123:~/electroDunas$ ls -al data/
total 11068
drwxrwxr-x 2 ubuntu ubuntu 4096 Oct 29 18:29 .
drwxrwxr-x 5 ubuntu ubuntu 4096 Oct 29 18:22 ..
-rw-rw-r-- 1 ubuntu ubuntu 591 Oct 29 18:29 .gitignore
-rw-rw-r-- 1 ubuntu ubuntu 6993 Oct 29 18:26 DATOSCLIENTE1.csv
-rw-rw-r-- 1 ubuntu ubuntu 97 Oct 29 18:26 DATOSCLIENTE1.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7000 Oct 29 18:28 DATOSCLIENTE10.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:28 DATOSCLIENTE10.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7000 Oct 29 18:29 DATOSCLIENTE11.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE11.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 1109179 Oct 29 18:29 DATOSCLIENTE12.csv
-rw-rw-r-- 1 ubuntu ubuntu 101 Oct 29 18:29 DATOSCLIENTE12.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7009 Oct 29 18:29 DATOSCLIENTE13.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE13.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7000 Oct 29 18:29 DATOSCLIENTE14.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE14.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7000 Oct 29 18:29 DATOSCLIENTE15.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE15.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7000 Oct 29 18:29 DATOSCLIENTE16.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE16.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 6999 Oct 29 18:29 DATOSCLIENTE17.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE17.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7000 Oct 29 18:29 DATOSCLIENTE18.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE18.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7000 Oct 29 18:29 DATOSCLIENTE19.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE19.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 6993 Oct 29 18:28 DATOSCLIENTE2.csv
-rw-rw-r-- 1 ubuntu ubuntu 97 Oct 29 18:28 DATOSCLIENTE2.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 7000 Oct 29 18:29 DATOSCLIENTE20.csv
-rw-rw-r-- 1 ubuntu ubuntu 98 Oct 29 18:29 DATOSCLIENTE20.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 855010 Oct 29 18:29 DATOSCLIENTE21.csv
-rw-rw-r-- 1 ubuntu ubuntu 100 Oct 29 18:29 DATOSCLIENTE21.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 874502 Oct 29 18:29 DATOSCLIENTE22.csv
-rw-rw-r-- 1 ubuntu ubuntu 100 Oct 29 18:29 DATOSCLIENTE22.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 873452 Oct 29 18:29 DATOSCLIENTE23.csv
-rw-rw-r-- 1 ubuntu ubuntu 100 Oct 29 18:29 DATOSCLIENTE23.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 858056 Oct 29 18:29 DATOSCLIENTE24.csv
-rw-rw-r-- 1 ubuntu ubuntu 100 Oct 29 18:29 DATOSCLIENTE24.csv.dvc
-rw-rw-r-- 1 ubuntu ubuntu 874043 Oct 29 18:29 DATOSCLIENTE25.csv
-rw-rw-r-- 1 ubuntu ubuntu 100 Oct 29 18:29 DATOSCLIENTE25.csv.dvc
```

Imagen que muestra cada uno de los archivos de datos crudos en formato CSV y el archivo de control de DVC (csv.dvc).



```
(env-dvc) ubuntu@ip-172-31-24-123:~/electroDunas$ git commit -m "Carga inicial de los datos de medicion para 30 client
as no regulados de ElectroDunas"
[main 68e615f] Carga inicial de los datos de medicion para 30 clientes no regulados de ElectroDunas
Committer: Ubuntu <ubuntu@ip-172-31-24-123.ec2.internal>
Your name and email address were configured automatically based
on your username and hostname. Please check that they are accurate.
You can suppress this message by setting them explicitly. Run the
following command and follow the instructions in your editor to edit
your configuration file:

    git config --global --edit

After doing this, you may fix the identity used for this commit with:

    git commit --amend --reset-author

31 files changed, 180 insertions(+)
create mode 100644 data/.gitignore
create mode 100644 data/DATOSCLIENTE1.csv.dvc
create mode 100644 data/DATOSCLIENTE10.csv.dvc
create mode 100644 data/DATOSCLIENTE11.csv.dvc
create mode 100644 data/DATOSCLIENTE12.csv.dvc
create mode 100644 data/DATOSCLIENTE13.csv.dvc
create mode 100644 data/DATOSCLIENTE14.csv.dvc
create mode 100644 data/DATOSCLIENTE15.csv.dvc
create mode 100644 data/DATOSCLIENTE16.csv.dvc
create mode 100644 data/DATOSCLIENTE17.csv.dvc
create mode 100644 data/DATOSCLIENTE18.csv.dvc
create mode 100644 data/DATOSCLIENTE19.csv.dvc
create mode 100644 data/DATOSCLIENTE2.csv.dvc
```

**Imagen que presenta el commit inicial, en donde son confirmados en el repositorio los 30 archivos de datos crudos del proyecto Electro dunas.**

```
(env-dvc) ubuntu@ip-172-31-24-123:~/electroDunas/data$ git commit -m "Cambio estructura"
[development e020ddb] Cambio estructura
Committer: Ubuntu <ubuntu@ip-172-31-24-123.ec2.internal>
Your name and email address were configured automatically based
on your username and hostname. Please check that they are accurate.
You can suppress this message by setting them explicitly. Run the
following command and follow the instructions in your editor to edit
your configuration file:

    git config --global --edit

After doing this, you may fix the identity used for this commit with:

    git commit --amend --reset-author

60 files changed, 30 insertions(+)
create mode 100644 data/raw/DATOSCLIENTE1.csv
rename data/{ => raw}/DATOSCLIENTE1.csv.dvc (100%)
create mode 100644 data/raw/DATOSCLIENTE10.csv
rename data/{ => raw}/DATOSCLIENTE10.csv.dvc (100%)
create mode 100644 data/raw/DATOSCLIENTE11.csv
rename data/{ => raw}/DATOSCLIENTE11.csv.dvc (100%)
create mode 100644 data/raw/DATOSCLIENTE12.csv
rename data/{ => raw}/DATOSCLIENTE12.csv.dvc (100%)
create mode 100644 data/raw/DATOSCLIENTE13.csv
rename data/{ => raw}/DATOSCLIENTE13.csv.dvc (100%)
create mode 100644 data/raw/DATOSCLIENTE14.csv
rename data/{ => raw}/DATOSCLIENTE14.csv.dvc (100%)
create mode 100644 data/raw/DATOSCLIENTE15.csv
rename data/{ => raw}/DATOSCLIENTE15.csv.dvc (100%)
create mode 100644 data/raw/DATOSCLIENTE16.csv
rename data/{ => raw}/DATOSCLIENTE16.csv.dvc (100%)
```

**Imagen que presenta el commit con el cambio de estructura, donde se mueven los archivos de la carpeta data a data/raw**

- *Reporte de trabajo en equipo de máximo 1 página.*

| <b>Actividad</b>                      | <b>Responsable</b>                              |
|---------------------------------------|-------------------------------------------------|
| <i>Definiciones</i>                   | <i>Luisa María – Juan Antonio – Juan Carlos</i> |
| <i>Creación Documento</i>             | <i>Juan Antonio</i>                             |
| <i>Creación Repositorio</i>           | <i>Juan Carlos</i>                              |
| <i>Cargue de los datos</i>            | <i>Juan Carlos</i>                              |
| <i>Creación maqueta Visualización</i> | <i>Luisa María</i>                              |
| <i>Exploración de los datos</i>       | <i>Juan Carlos</i>                              |
| <i>Verificación del Documento</i>     | <i>Luisa María – Juan Antonio – Juan Carlos</i> |



**Bibliografía:**

- <https://www.celsia.com/en/blog-celsia/conoce-todo-lo-que-debes-saber-sobre-energia-reactiva-por-que-se-refleja-en-tu-factura-la-energia-reactiva/>
- <https://www.plena-energia.com/post/energia-activa-reactiva#:~:text=En%20pocas%20palabras%2C%20la%20activa,se%20le%20llama%20energ%C3%ADa%20reactiva.>
- [https://es.wikipedia.org/wiki/Tensi%C3%B3n\\_\(electricidad\)](https://es.wikipedia.org/wiki/Tensi%C3%B3n_(electricidad)).
- Materiales del curso Gerencia de Proyectos de Analítica. Caso Electro Dunas.