

The background image shows a panoramic view of a residential area in a cold, possibly northern European town. The houses are built on a hillside, with their roofs and ground covered in a thick layer of white snow. The houses are painted in various colors, including white, blue, yellow, red, and grey. Some have traditional gabled roofs, while others are modern. Bare trees stand between the houses, and a few evergreen trees are scattered throughout the scene.

Proforecast Estonia

Modelo predictivo para la autogestión energética en Estonia

Energía solar: Potencia Global en Ascenso

- ▶ En 2030, la demanda eléctrica anual deberá ser 100% cubierta por energías renovables¹
- ▶ El objetivo de la neutralidad de la huella de Carbono, para 2050
- ▶ La guerra en Ucrania³, un motivo de peso para la autosuficiencia energética.
- ▶ El balance energético, la última de las claves



El caso en estudio, Estonia

- ▶ Su situación geográfica la hace especialmente sensible a la Guerra de Ucrania.
- ▶ Ya ha comenzado a tomar medidas para alcanzar los acuerdos de sostenibilidad



Estonia acelera el cambio a la energía verde, adoptando las energías renovables en lugar del esquisto bituminoso

by Noticias ASG • 20 de noviembre.

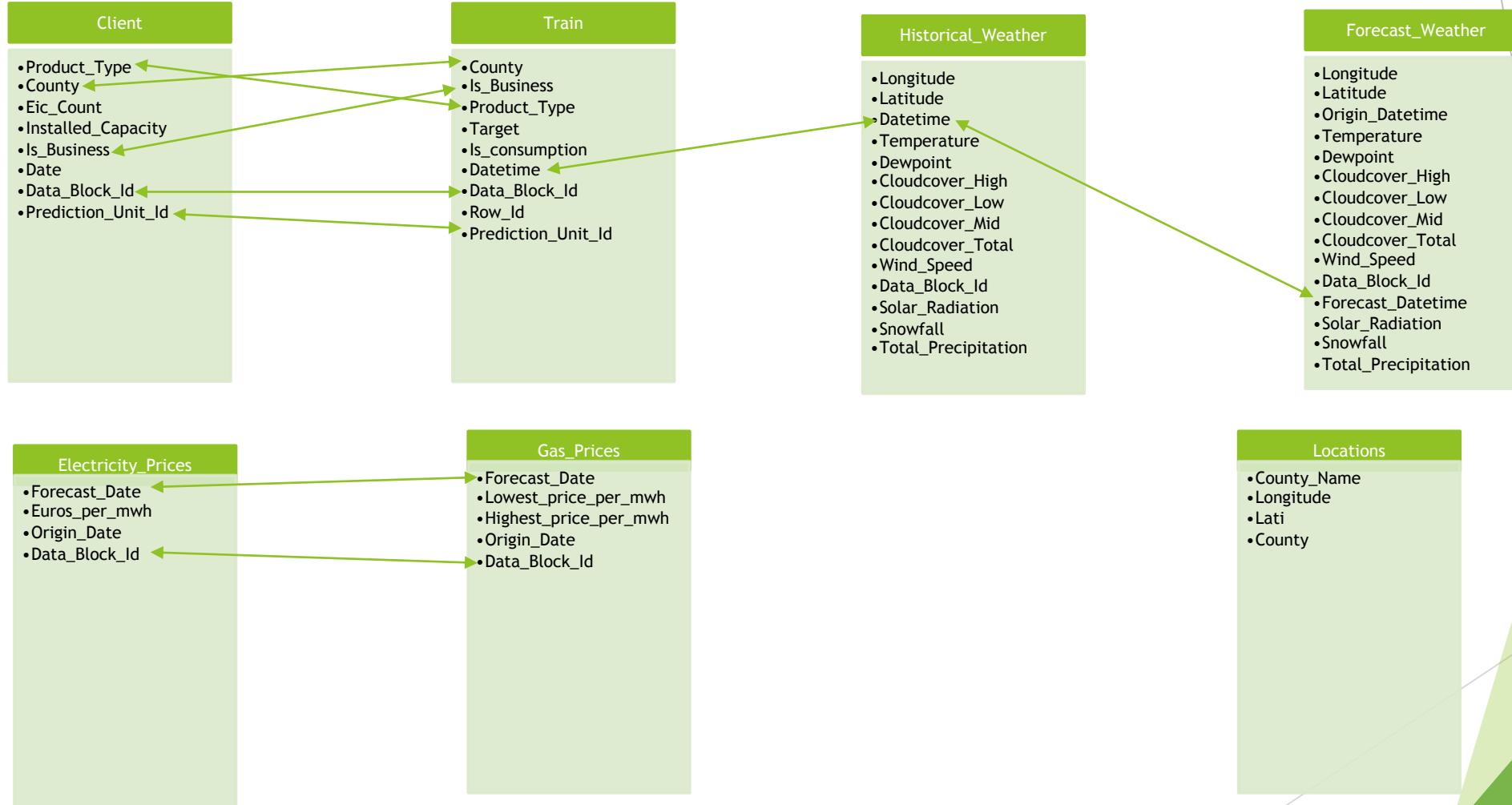
Compartir: [f](#) [t](#) [in](#)

The background image shows a large industrial complex, likely a power plant or refinery, with multiple tall smokestacks emitting plumes of smoke. A large sign on one of the buildings prominently displays the text "Estrella Energía". The sky is overcast and hazy.

Panorama del problema: Fuente de los datos

Panorama de datos:

Relaciones entre tablas

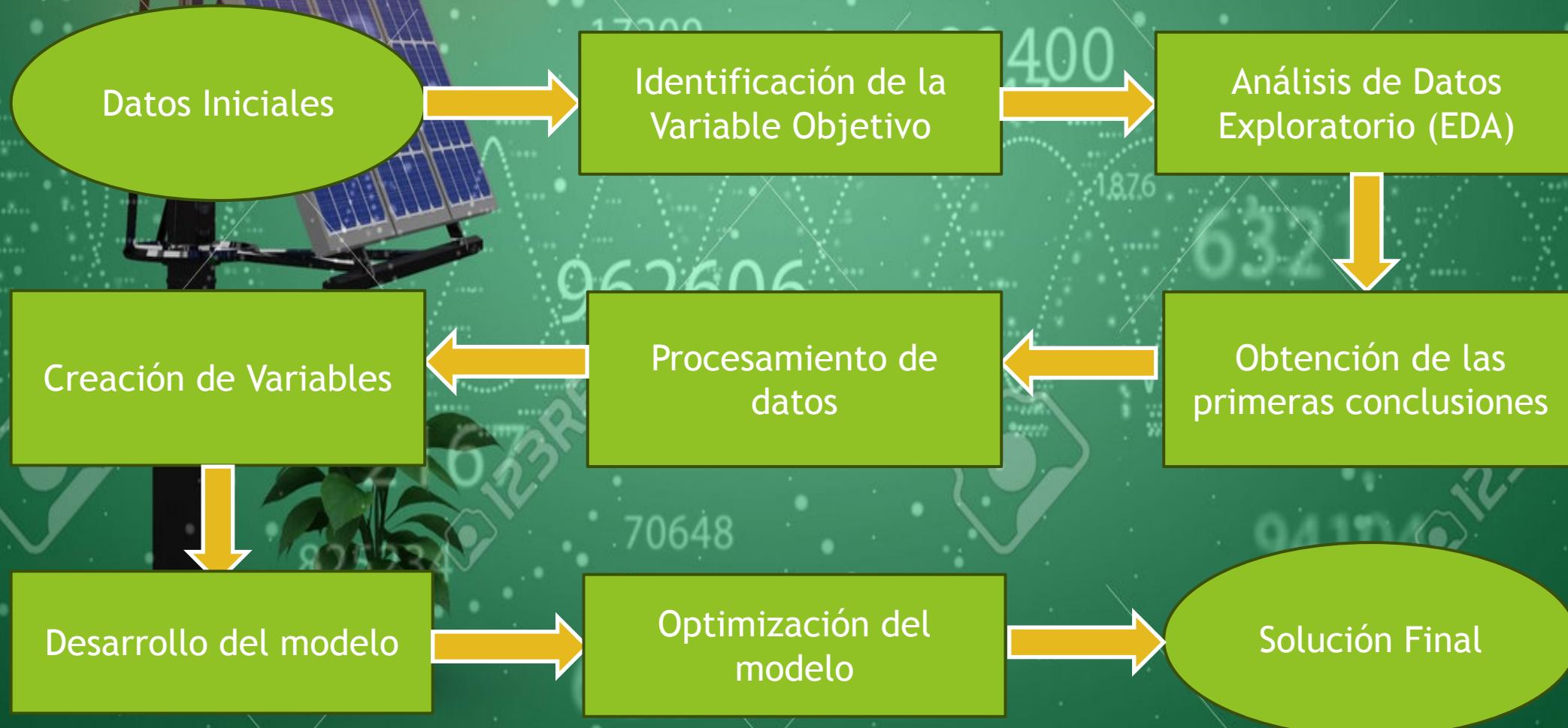


Resumen de datos: Valores más destacables

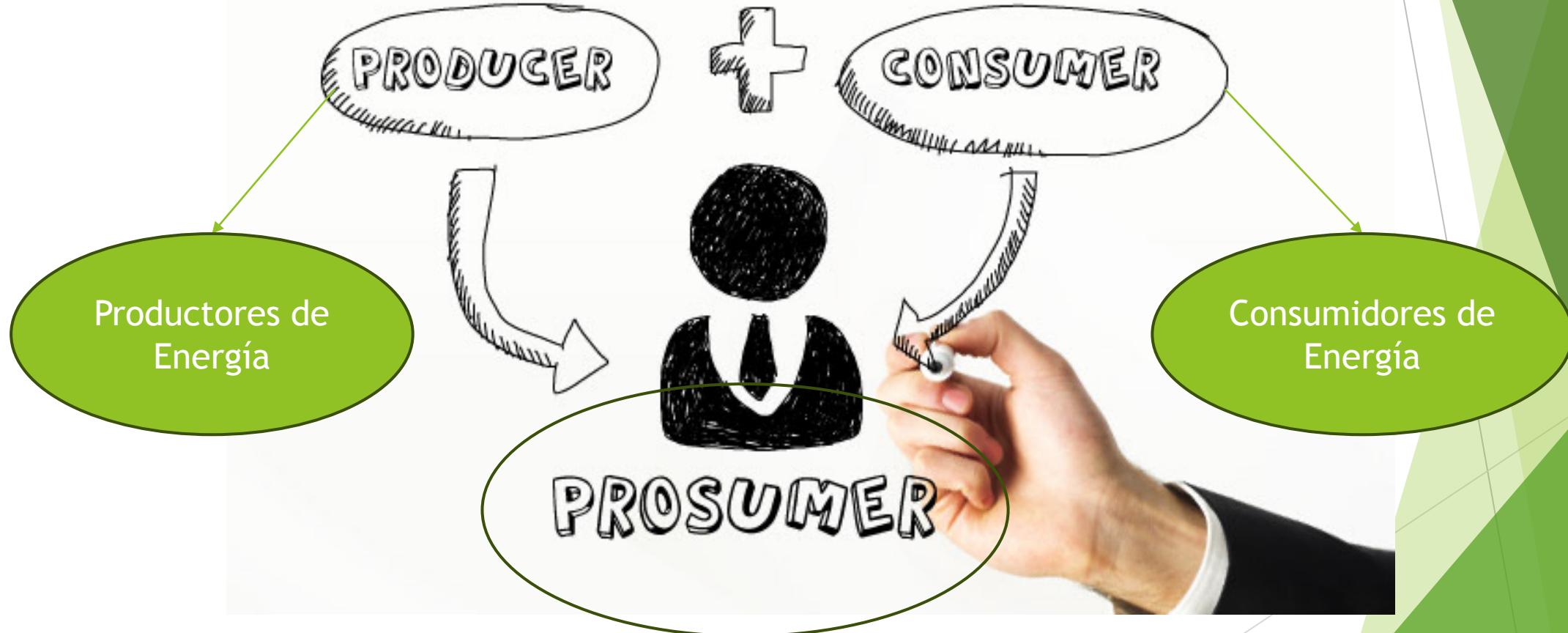
Resmuen de los datos	
Tiempo del muestreo	3 años
Número total de prosumidores	69 usuarios
Tipos de contrato	4
Tipos de usuario	2
Muestreo por día	1 medida/hora
Total medidas climáticas	14

Estructurando la solución:

Diagrama de flujo



Objetivos detallados: Definiendo el éxito

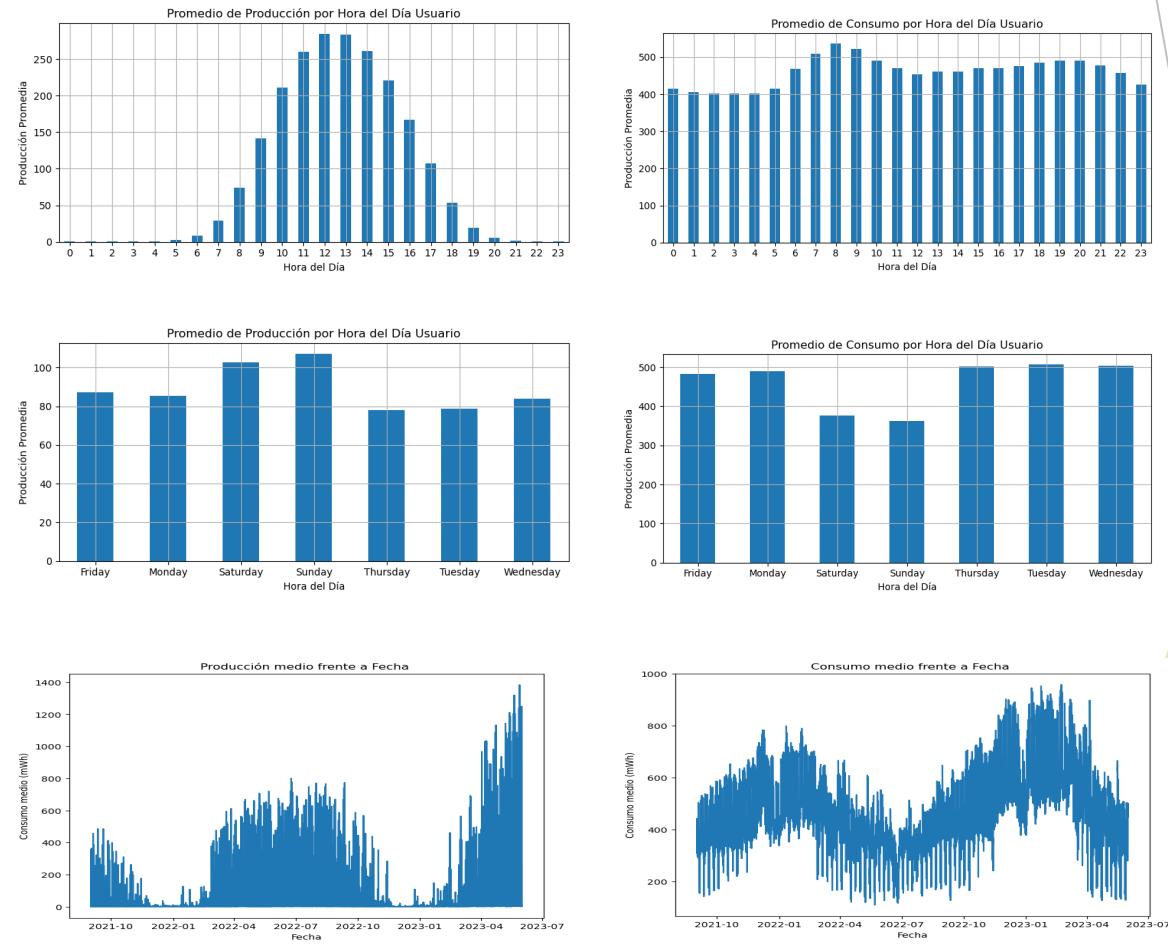
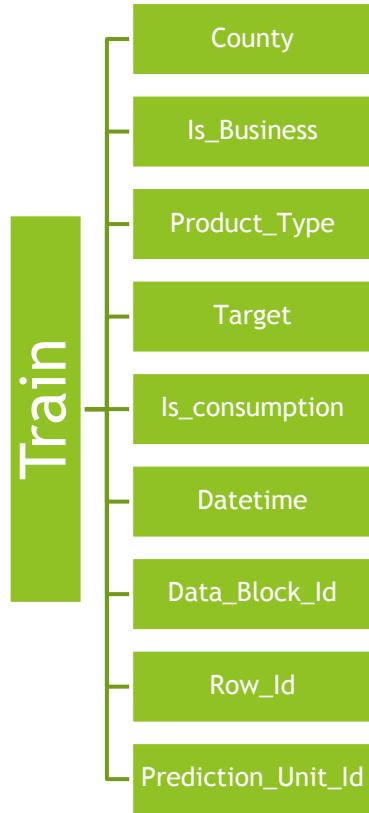


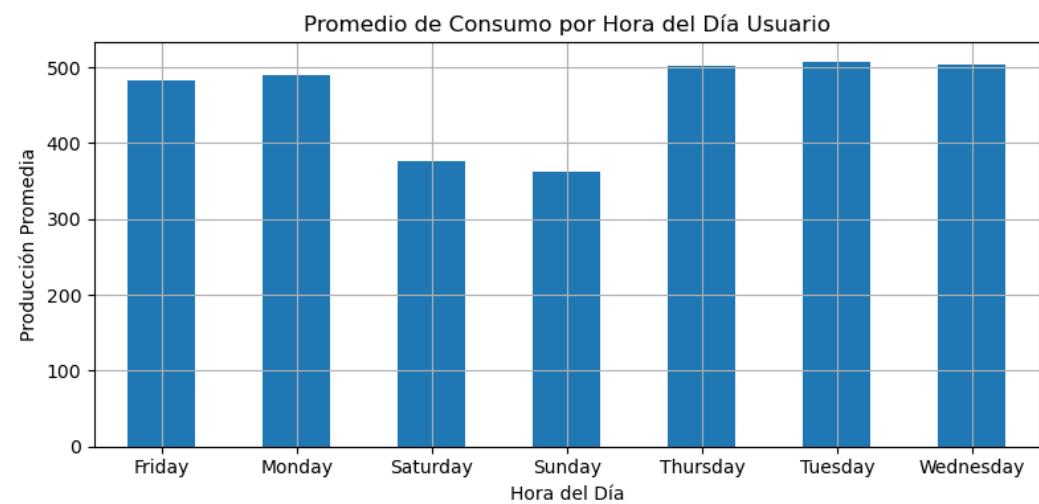
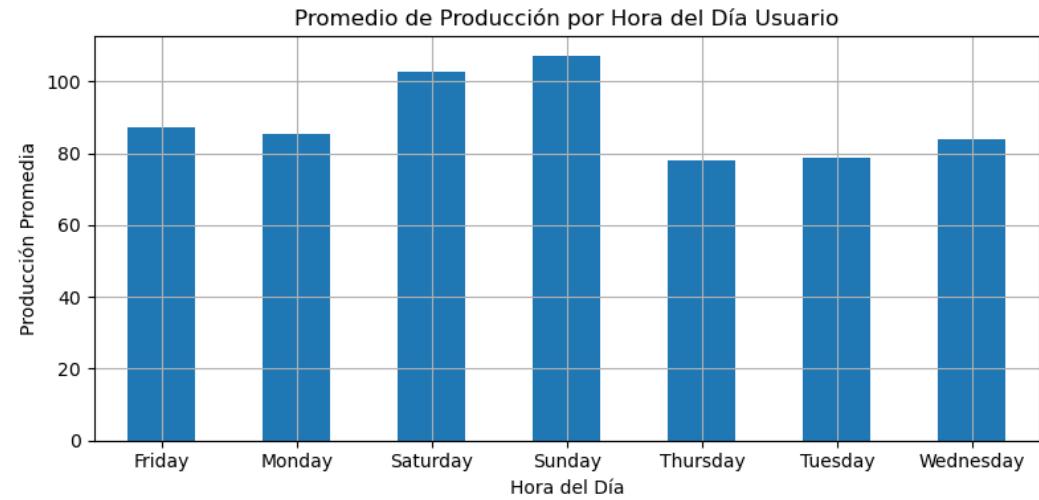


Despliegue del proyecto: Análisis y preparación de datos

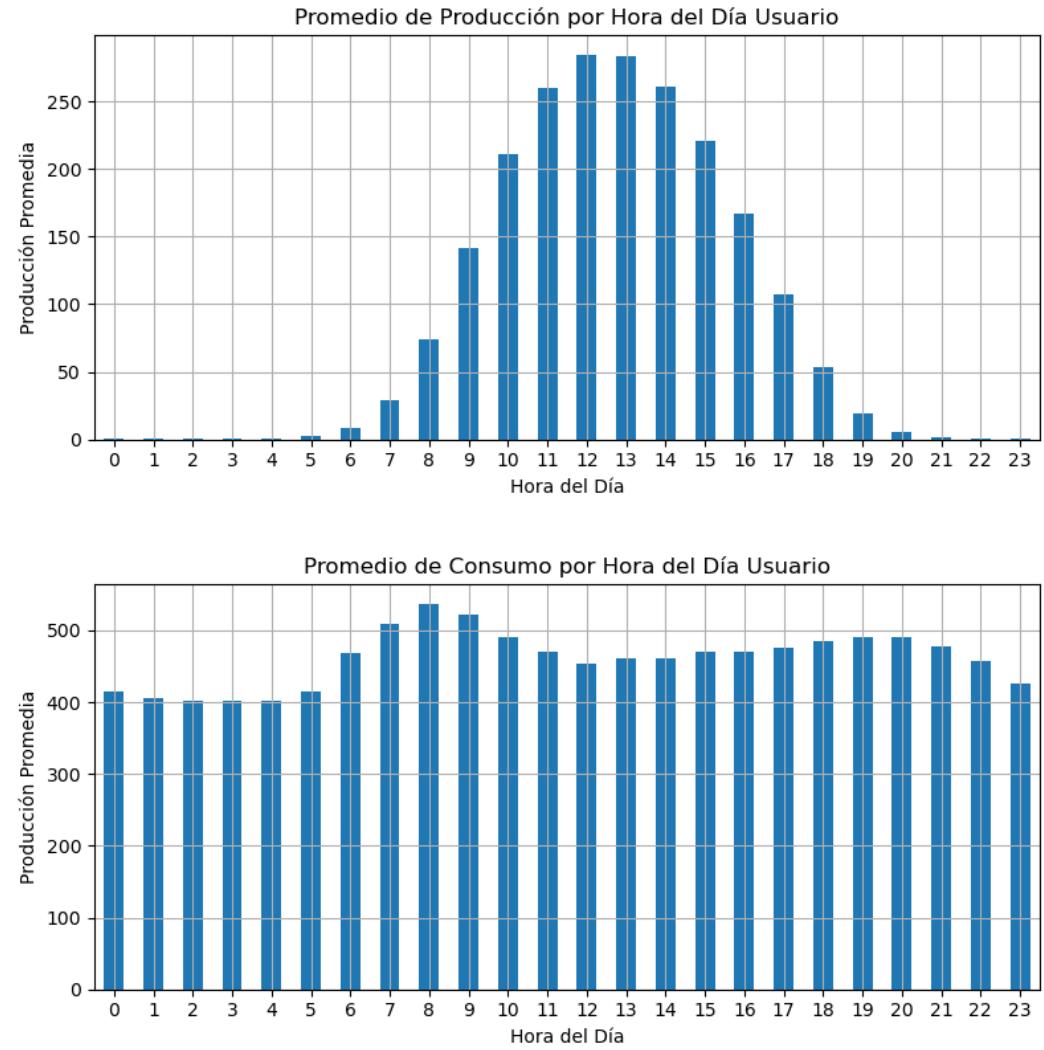
EDA:

Análisis Exploratorio de los Datos

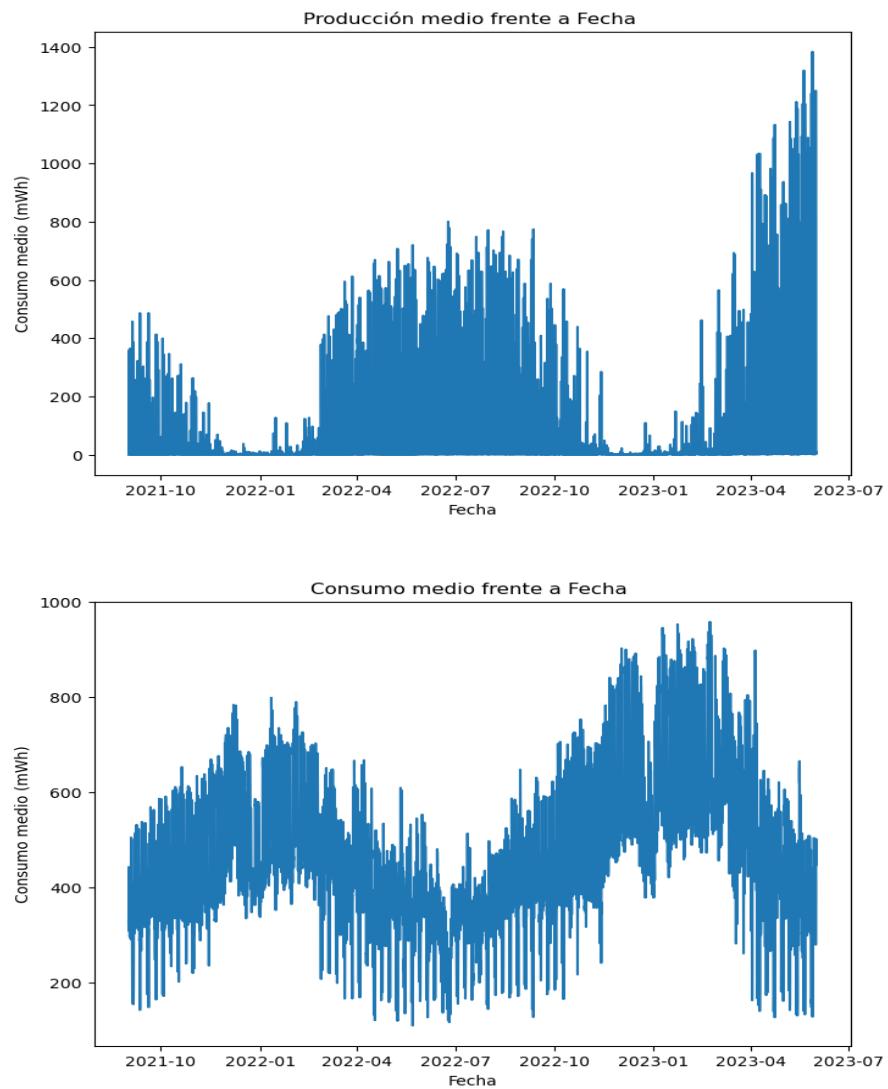




Promedio de
Producción
y Consumo
por Día



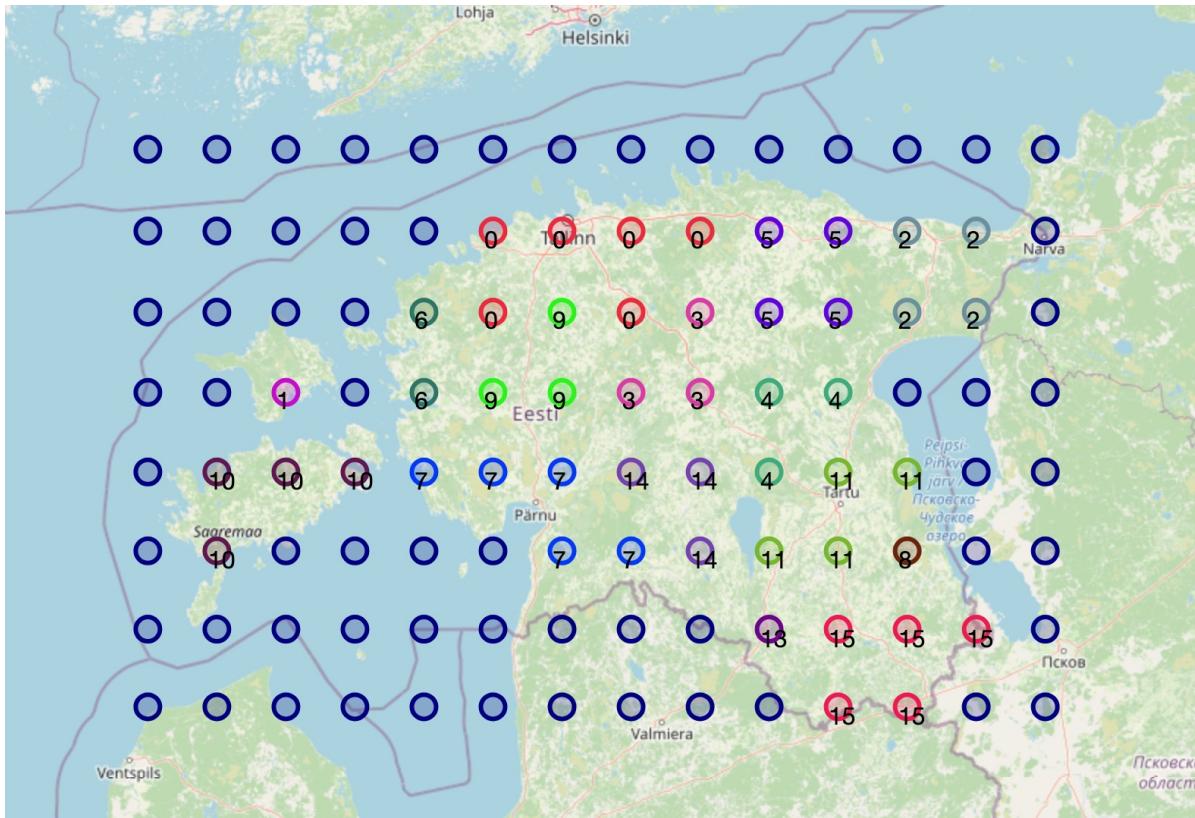
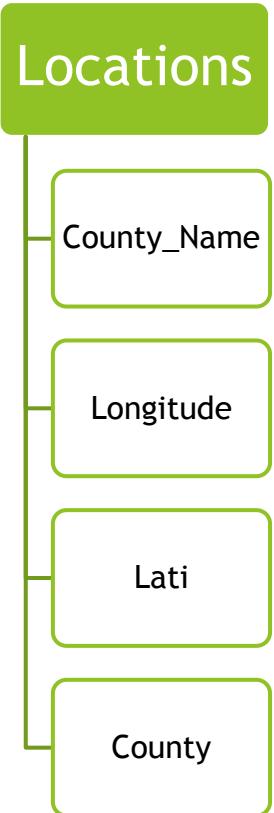
Promedio de Producción y Consumo por Hora



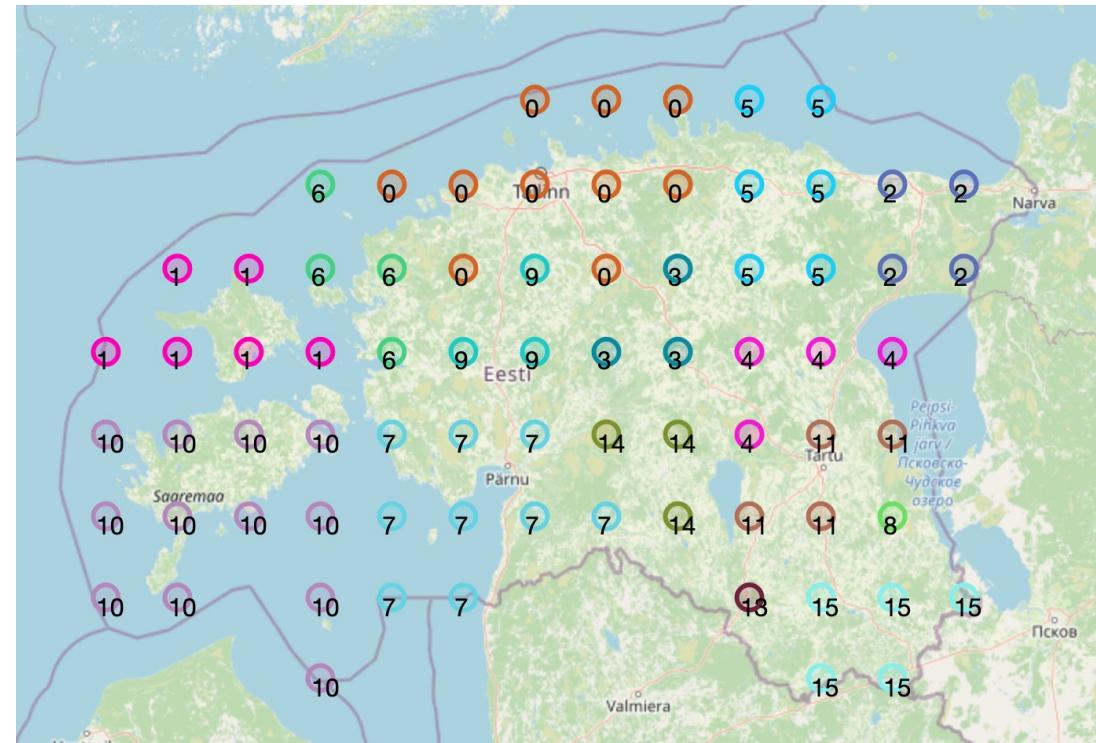
Promedio de
Producción y
Consumo a lo
largo del año

EDA:

Análisis Exploratorio de los Datos



► Mapeado de los condados de Estonia una vez filtrados respecto a las fronteras



Descubrimientos claves y conclusiones preliminares

1. Comportamiento de la producción y el consumo → Estacionalidad
2. Separar variables climáticas atendiendo a qué target queremos predecir.
3. Clusterización de las variables climáticas → Ayudará a la predicción del target
4. Se tomarán los valores climáticos medios atendiendo su cantón

Preparación y Limpieza de Datos

- ▶ 1er caso → Target: PRODUCCIÓN
 - ▶ Filtrado por variable "Is_Consumption"
 - ▶ Modificación del target → Se predecirá la eficiencia de la placa
 - ▶ Modificación de la variable Datetime
 - ▶ Unión de tablas de Clima y Train para conocer condiciones climáticas del cantón



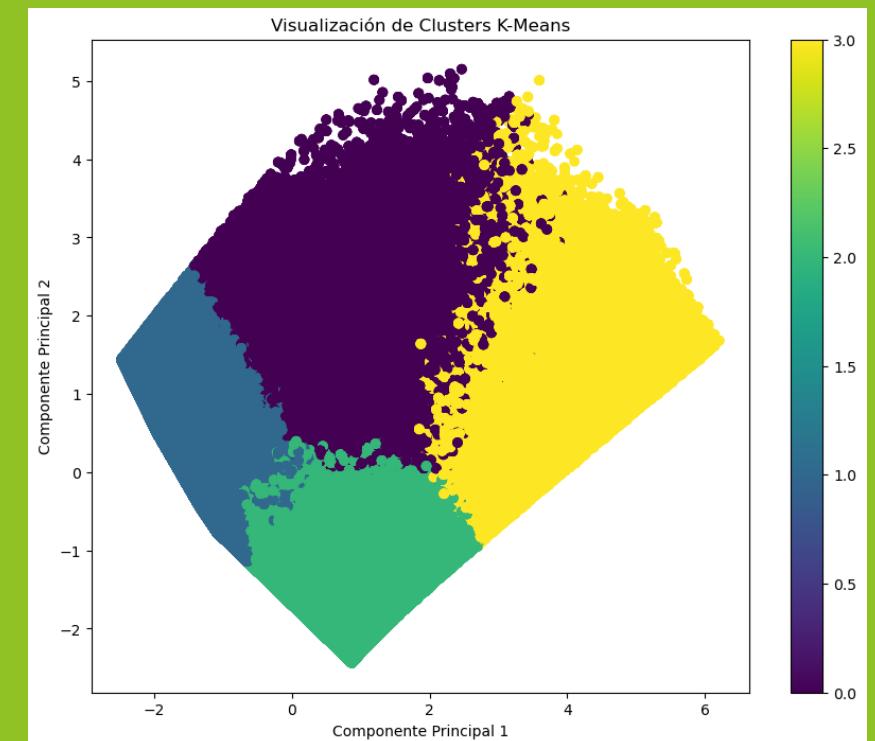
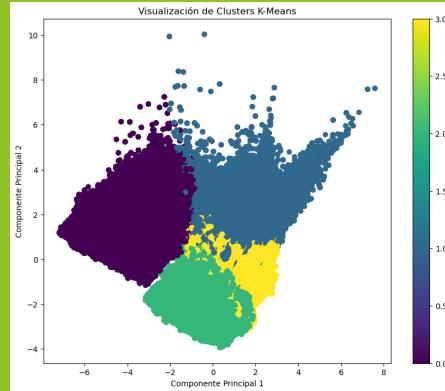
Preparación y Limpieza de Datos

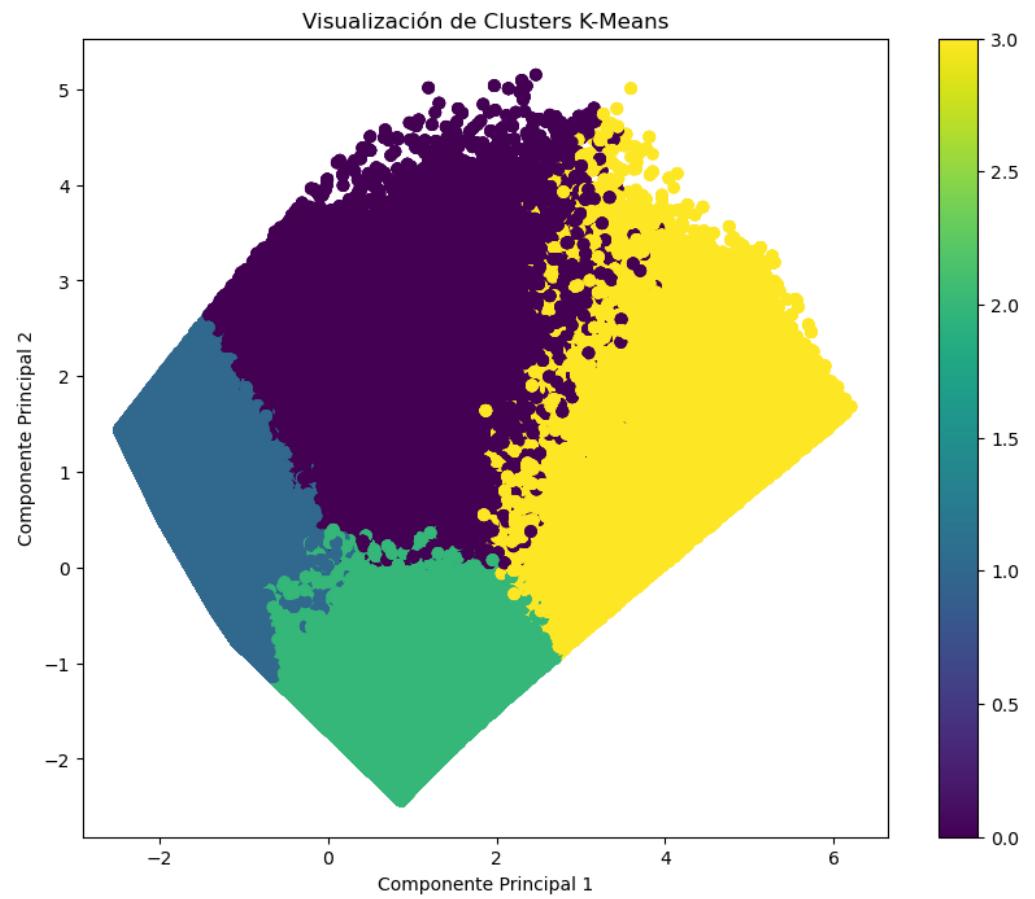
- ▶ 2o caso → Target: CONSUMO
 - ▶ Filtrado por variable "Is_Consumption"
 - ▶ No se modifican el target → Se predecirá directamente el consumo
 - ▶ Modificación de la variable Datetime
 - ▶ Unión de tablas de Clima y Train para conocer condiciones climáticas del cantón



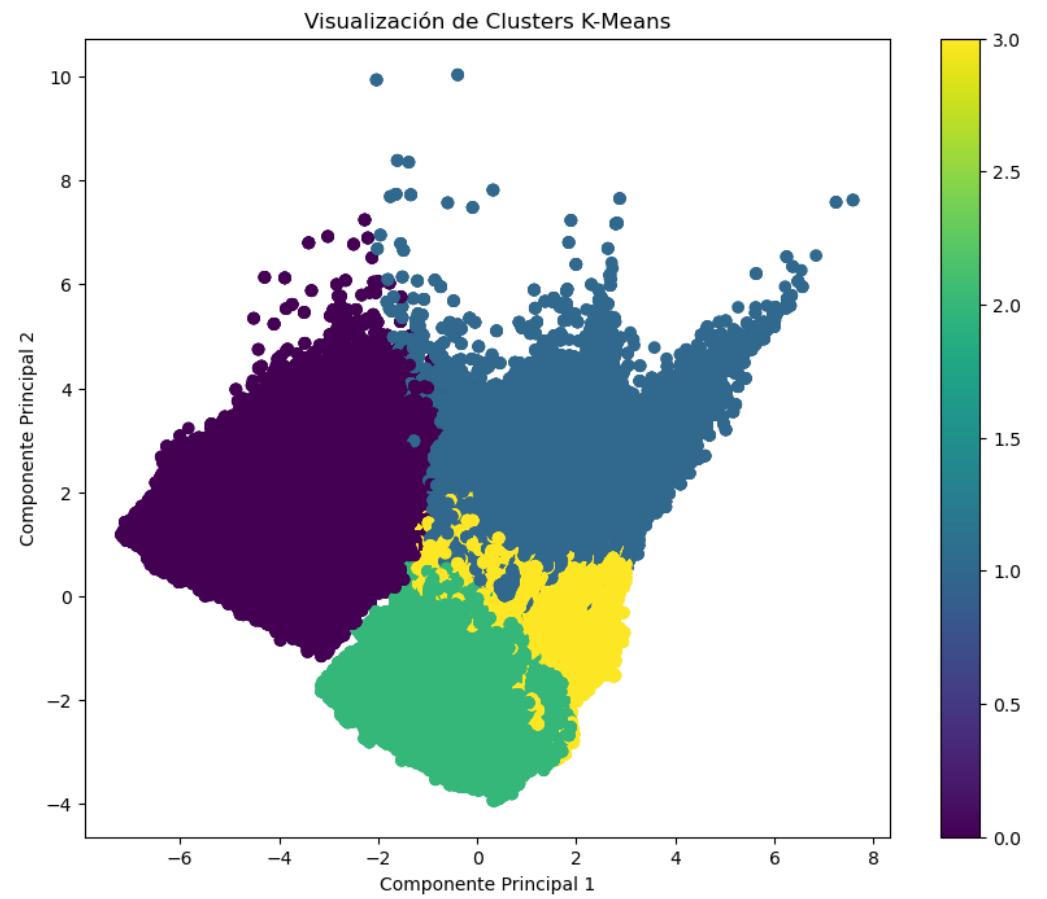
Ingeniería de Características: Construyendo variables para Mejor Predicción

- ▶ Creación de nuevas variables → Laggeo temporal.
- ▶ Otro concepto importante → Clusters Climáticos





Ingeniería de Características: Clusters Climáticos para Producción



Ingeniería de Características: Clusters Climáticos para Consumo

Creación de variables

- ▶ La clave ha sido el laggeo.
- ▶ Se han hecho laggeos tanto para los clusters de los climas como para los consumos/producciones anteriores.



El laggeo como solución

Visualización de las variables creadas para la producción respecto a los DATOS de DÍAS ANTERIORES

(Usuario: 1, Hora: 12:00)						
Fecha	Producción (kWh)	Perf_Rate	Shift 1	Shift 2	Shift 3	Shift 4
10/06/2022	47.408	0.5070	0.5891	0.5886	0.6222	0.3353
11/06/2022	41.161	0.4402	0.5070	0.5891	0.5886	0.6222
12/06/2022	22.870	0.2446	0.4402	0.5070	0.5891	0.5886
13/06/2022	67.472	0.7216	0.2446	0.4402	0.5070	0.5891

El laggeo como solución

Visualización de las variables creadas para la producción respect a la
CLUSTERIZACIÓN CLIMÁTICA

(Usuario: 1, Hora: 12:00, Cluster: 1)						
Fecha	Producción (kWh)	Perf_Rate	Shift 1	Shift 2	Shift 3	Shift 4
19/03/2023	19.594	0.2227	0.2023	0.2126	0.2725	0.2654
22/03/2023	21.446	0.2437	0.2227	0.2023	0.2126	0.2725
26/03/2023	24.086	0.2737	0.2437	0.2227	0.2023	0.2126
27/03/2023	20.698	0.2352	0.2737	0.2437	0.2227	0.2023

Estrategias de Modelado

Hiperparámetros	Valor
N_estimators	100
Max_Depth	30
Min_samples_Split	5

- ▶ Se emplearon dos para la regresión → XGBoost y RandomForestRegressor
- ▶ Método de validación → 80% de train, 20% de test
- ▶ Finalmente, se escoge RandomForestRegressor:
 - ▶ Menos propenso al sobreajuste
 - ▶ Se obtienen mejores resultados en las métricas seleccionadas
- ▶ Hiperparámetros seleccionados:

Resultados obtenidos

TRAIN				
Producción	MAE	MSE	RMSE	R2
RFR	12.87	3762.60	61.34	0.95

TEST				
Producción	MAE	MSE	RMSE	R2
RFR	16.78	5892.81	76.76	0.92

TRAIN				
Consumo	MAE	MSE	RMSE	R2
RFR	13.02	4004.36	63.28	0.96

TEST				
Consumo	MAE	MSE	RMSE	R2
RFR	16.41	6050	77.78	0.93

Conclusiones



Se puede afirmar la hipótesis de que los targets pasados son útiles a la hora de poder obtener predecir los targets futuros.



Clusterizar el clima sirve para poder acceder a valores anteriores donde las condiciones atmosféricas también fueron similares.



Los valores de precio de gas y electricidad no presentan un peso real en el modelo realizado.

Trabajo a Futuro

- ▶ Utilizar los valores climáticos de predicciones → poder predecir con tiempo el consumo
- ▶ Mejora del overfitting presente en ambos modelos.
- ▶ Estudiar la posibilidad de aumentar el número de clusters → “necesidad” de más datos
- ▶ Estudio de la creación de nuevas variables



¡MUCHAS GRAIAS!

Si tenéis alguna duda, no dudéis en preguntar!!



Nicolás Gil Grandal

Bibliografía

¹IEA (2023), Estonia 2023, IEA, Paris
<https://www.iea.org/reports/estonia-2023>, Licence: CC BY 4.0

²<https://www.iea.org/reports/renewables-2023>

³<https://es.euronews.com/2022/09/02/rusia-cierra-el-grifo-del-gas-a-europa-interrumpe-el-suministro-de-forma-indefinida>

