

# Caso: Preparación y Análisis de Datos de Precios de Combustibles en Argentina

Recientemente usted ha sido incorporado a la Secretaría de Energía, justo después de haber culminado con éxito la materia Analítica de Datos en la carrera de Negocios Digitales en la Universidad de San Andrés. Esta experiencia no solo lo armó con un conocimiento profundo en análisis de datos, sino que también agudizó su habilidad para enfrentar y resolver problemas complejos mediante la utilización de herramientas avanzadas de análisis y modelado de datos. Con una energía renovada y técnicas de vanguardia, se prepara para aportar una nueva perspectiva al equipo de Análisis de Datos de la Secretaría.

Se le asigna una primera tarea: analizar los precios de combustibles en la Argentina. Reportando directamente al Director, está ansioso por zambullirse en los desafíos que presenta el análisis de los precios de combustibles en el país. Su misión es clara: **identificar patrones significativos** brindando materia prima para que la Secretaría pueda ofrecer recomendaciones estratégicas que estén bien fundamentadas en los datos recabados.

Su capacidad única para desentrañar complejidades técnicas y transformarlas en insights claros y accionables lo convierte en un activo invaluable para la dirección.

## Contexto:

En el contexto de incrementos sostenidos en los precios del combustible en Argentina, la Secretaría de Energía busca **comprender en profundidad las dinámicas de precios a nivel nacional**. Este análisis es crucial para diseñar políticas que aseguren la equidad y la eficiencia en el acceso al combustible. En este escenario, usted ha sido contratado como analista de datos para llevar a cabo una investigación exhaustiva sobre los precios de los combustibles en las estaciones de servicio de todo el país.

## Desafío:

Los precios del combustible son un componente vital de la economía, afectando desde el coste de vida hasta la viabilidad de las empresas. Su tarea como analista de datos es navegar a través de un conjunto complejo de datos, comprendiendo las tendencias, realizando visualizaciones exploratorias, y abordando problemas de calidad de datos. Se espera que, mediante su análisis, pueda ofrecer una visión clara de la situación actual de los precios de los combustibles, identificar patrones, anomalías y posibles causas de las fluctuaciones de precios.

## Objetivos Específicos:

1. Entender el Conjunto de Datos: Familiarizarse con la estructura, los tipos de datos y la cobertura geográfica del conjunto de datos proporcionado por la Secretaría de Energía.
2. Hacer un análisis exploratorio de datos (EDA)

3. Ejecutar el siguiente script y observar los resultados. Agregar, por lo menos, dos visualizaciones al análisis:

<https://colab.research.google.com/drive/1QbgpwYqnXInFRiPAsBz149NoX9Ct1U1O#scrollTo=tdxyfojRKr1R>

4. Calidad de Datos: Identificar y resolver problemas de calidad en los datos, como valores faltantes, duplicados, o incoherencias.
5. Volver a ejecutar el script del punto 3 y analizar las diferencias entre el antes y el después

### Pautas:

El trabajo se realizará en grupos de 5 integrantes. La entrega será el día lunes 1 de septiembre por el campus y el martes 2 de septiembre los equipos presentarán los resultados.

### Metadata del archivo

- indice\_tiempo: Fecha del registro, indicando el mes y año al que corresponde la información.
- idempresa: Identificador numérico único para cada empresa.
- cuit: CUIT de la empresa.
- empresa: Nombre de la empresa.
- direccion: Dirección de la estación de servicio.
- localidad: Localidad donde se encuentra la estación de servicio.
- provincia: Provincia de la estación de servicio.
- region: Región geográfica de la estación de servicio.
- idproducto: Identificador numérico único para cada tipo de producto (combustible).
- producto: Nombre del tipo de combustible.
- idtipohorario: Identificador numérico para el tipo de horario (diurno/nocturno).
- tipohorario: Tipo de horario en el que se reporta el precio (Diurno/Nocturno).
- precio: Precio del combustible en pesos argentinos.
- fecha\_vigencia: Fecha y hora exacta de la vigencia del precio reportado.
- idempresabandera: Identificador numérico único para la bandera de la empresa.
- empresabandera: Nombre de la bandera de la empresa.
- latitud: Latitud geográfica de la estación de servicio.
- longitud: Longitud geográfica de la estación de servicio.
- geojson: Representación en formato GeoJSON de la ubicación geográfica de la estación de servicio, incluyendo coordenadas de longitud y latitud.

### Entregables:

Un documento (formato pdf, máximo 5 carillas) que incluya una descripción del conjunto de datos, los problemas de calidad de datos detectados, sus correspondientes soluciones y los insights encontrados.