



Instituto Tecnológico
de Buenos Aires

TRATAMIENTO DE DATASET, ANÁLISIS Y GRÁFICOS

Federico Kaplun

—

Análisis exploratorio de precios vigentes en surtidor



Objetivo



POLÍTICA Y ECONOMÍA

ECONOMÍA

El Gobierno acordó postergar el aumento de la nafta: ¿Cuándo sube?

Luego de una reunión entre el Gobierno Nacional y representantes del sector de los combustibles se acordó que el aumento del 4% en la nafta se postergue



Mientras más chica es la localidad, menor es el acceso al gas natural

LOCALES 28 DE MARZO DE 2023 POR TRIBUNA




Contexto

En el 2017, se aprobó la resolución Resolución 314/2016, que le exige a todos los titulares de surtidores cargar datos de precios mensualmente.

Fuentes:

- [Guía de proveedores de estaciones de servicio](#)
- [Surtidores.com](#)

El Data Frame

 **Secretaría de Energía**

[Datasets](#) [Series](#) [Organizaciones](#) [SIG](#) [Visualizaciones](#) [APIs](#)

[Datos Energía](#) / [Dataset](#) / Recurso

Precios vigentes en surtidor - Resolución 314/2016

Precios vigentes de naftas, gasoil y GNC, según declaraciones realizadas por EESS en el marco de la Resolución 314/2016 del Ministerio de Energía y Minería, en AR\$/litro

[DESCARGAR](#)

Cantidad de registros: 36.936
Frecuencia de actualización: Cada hora

Fuente:

- [Secretaria de Energia](#)

El Data Frame

```
RangeIndex: 36942 entries, 0 to 36941
Data columns (total 19 columns):
#   Column              Non-Null Count  Dtype
---  ---
0   indice_tiempo        36942 non-null  object
1   idempresa            36942 non-null  int64
2   cuit                 36942 non-null  object
3   empresa              36942 non-null  object
4   direccion            36942 non-null  object
5   localidad            36942 non-null  object
6   provincia            36942 non-null  object
7   region              30553 non-null  object
8   idproducto           36942 non-null  int64
9   producto             36942 non-null  object
10  idtipohorario        36942 non-null  int64
11  tipohorario          36942 non-null  object
12  precio               36942 non-null  float64
13  fecha_vigencia       36942 non-null  object
14  idempresabandera     36942 non-null  int64
15  empresabandera       36942 non-null  object
16  latitud              36930 non-null  float64
17  longitud             36930 non-null  float64
18  geojson              36930 non-null  object
dtypes: float64(3), int64(4), object(12)
```

```
RangeIndex: 36942 entries, 0 to 36941
Data columns (total 12 columns):
#   Column              Non-Null Count  Dtype
---  ---
0   indice_tiempo        36942 non-null  datetime64[ns]
1   localidad            36942 non-null  object
2   provincia            36942 non-null  object
3   region              30553 non-null  object
4   producto             36942 non-null  object
5   tipohorario          36942 non-null  object
6   precio               36942 non-null  float64
7   fecha_vigencia       36942 non-null  object
8   empresabandera       36942 non-null  object
9   latitud              36930 non-null  float64
10  longitud             36930 non-null  float64
11  geojson              36930 non-null  object
dtypes: datetime64[ns](1), float64(3), object(8)
```

	precio
count	30551.000000
mean	210.829991
std	73.162275
min	7.000000
25%	191.400000
50%	220.100000
75%	256.100000
max	365.000000

Limpieza

Diagrama de caja antes de eliminar outliers

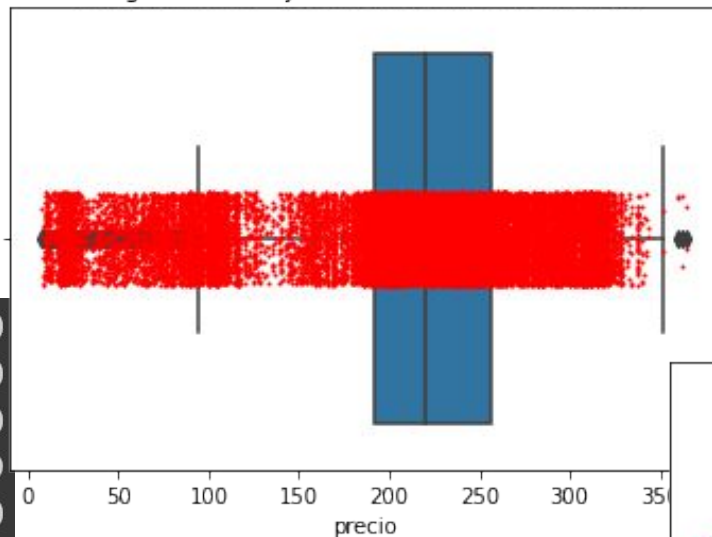
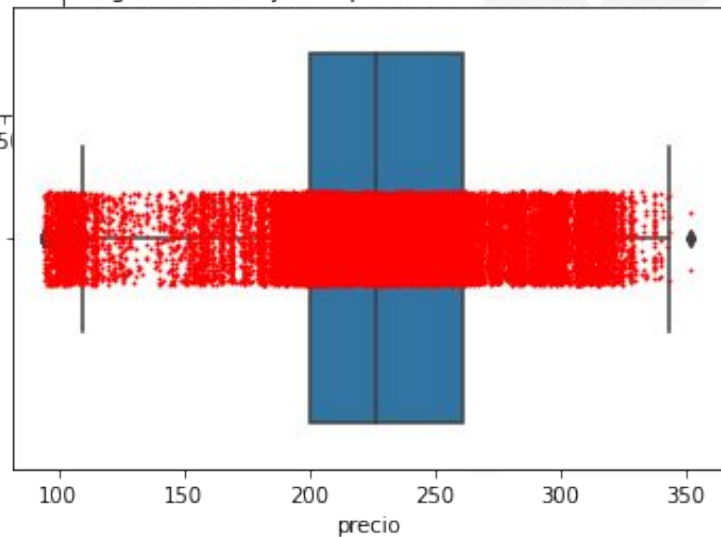
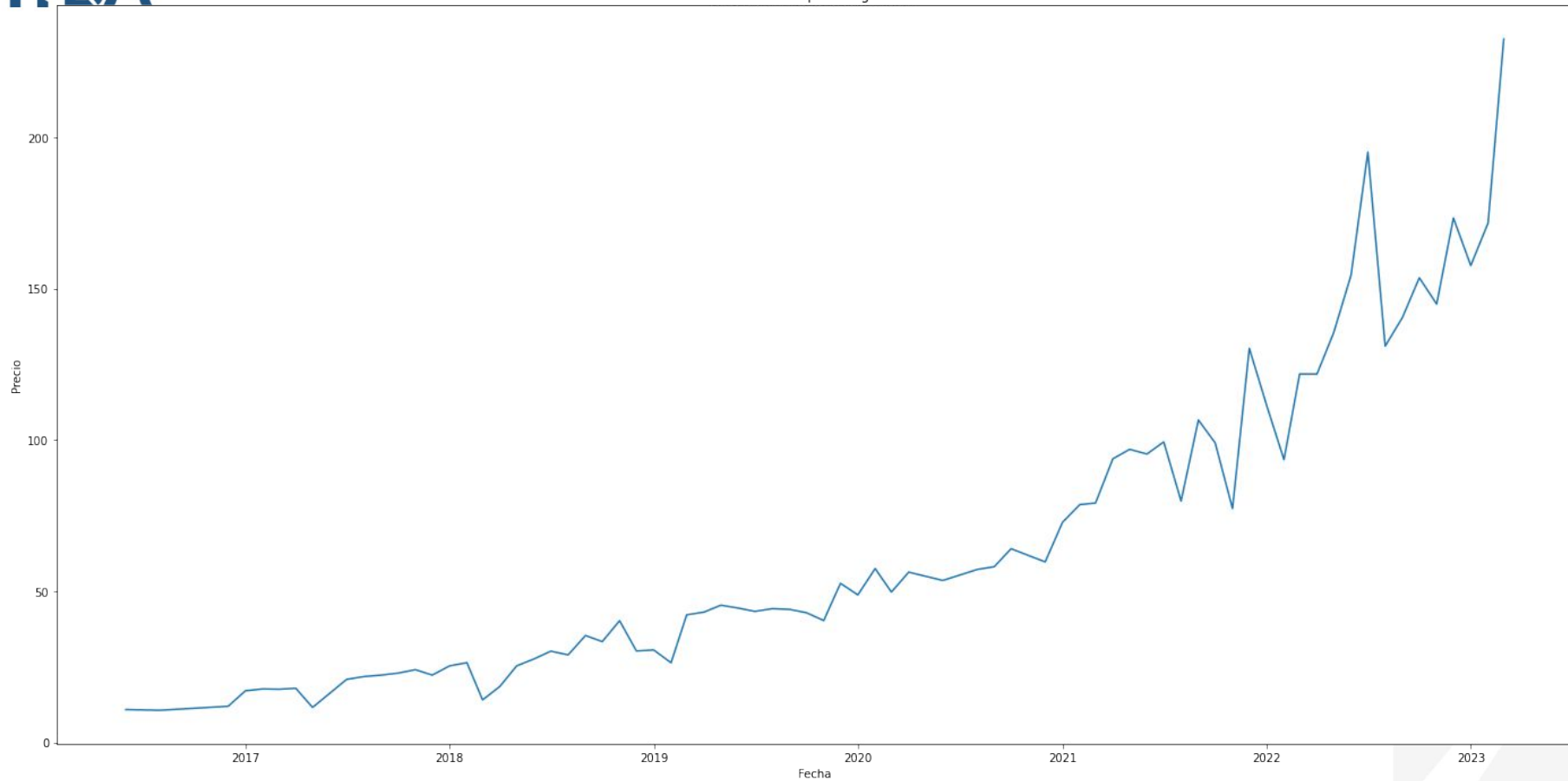
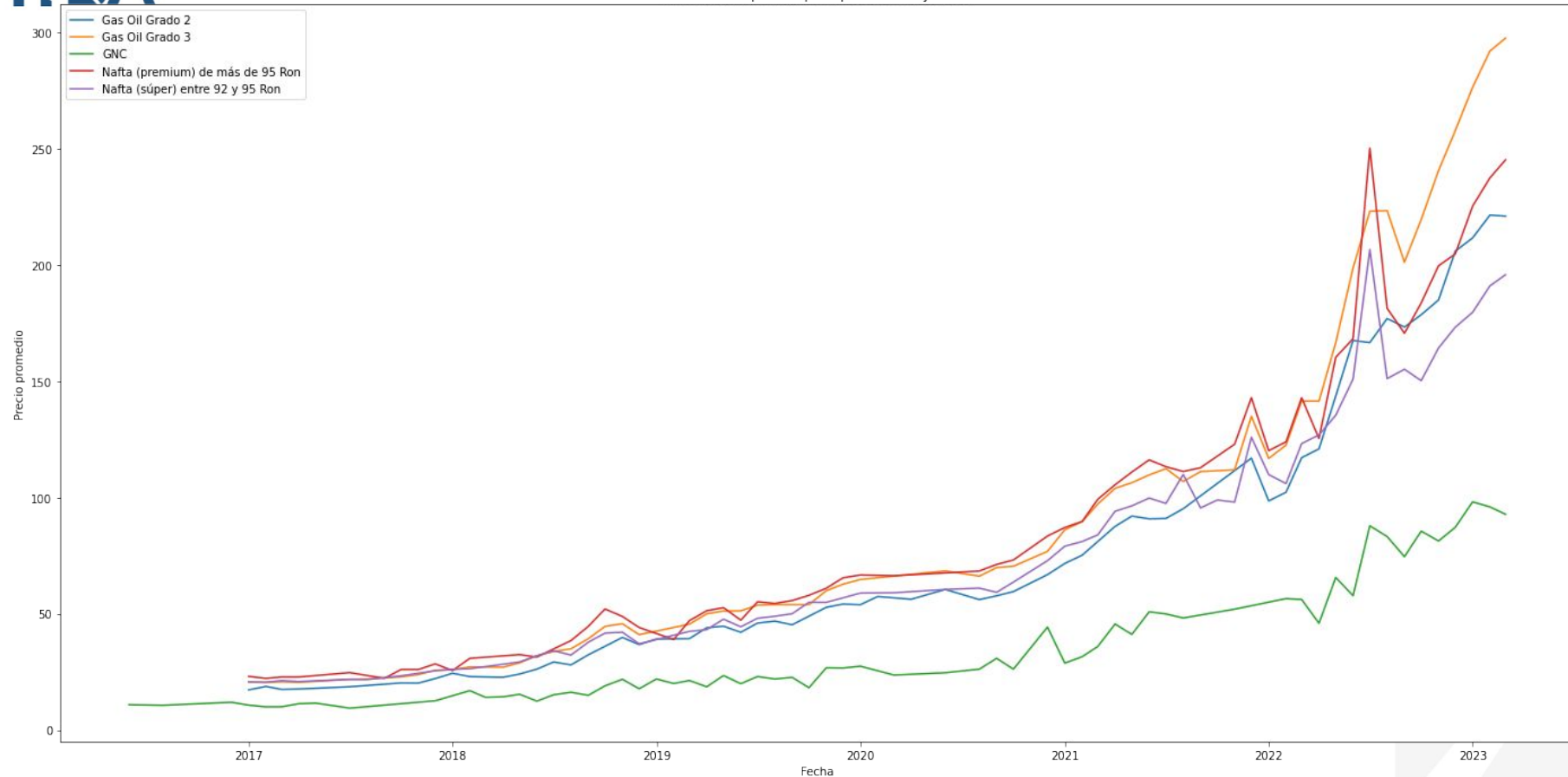


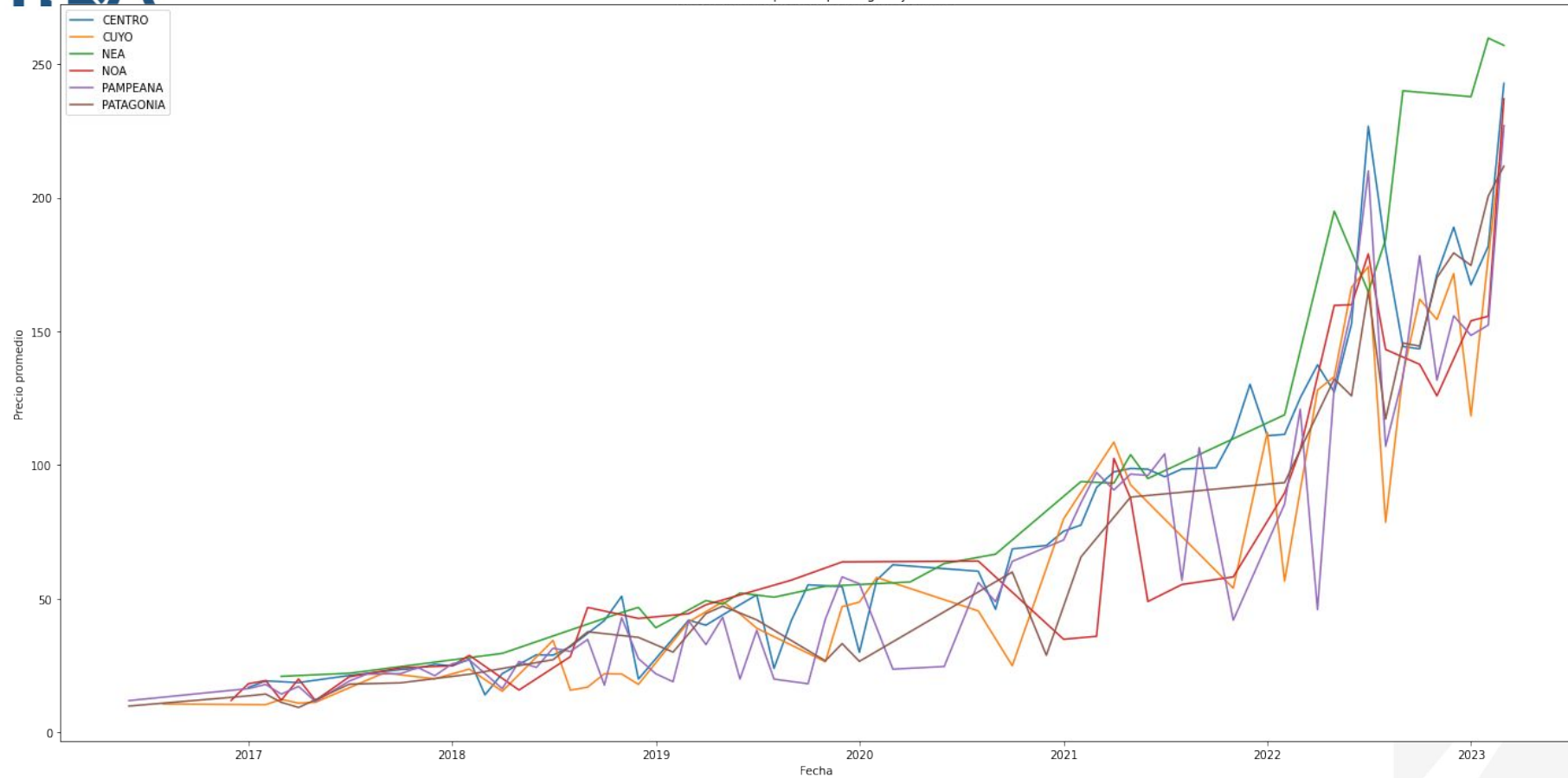
Diagrama de caja después de eliminar outliers

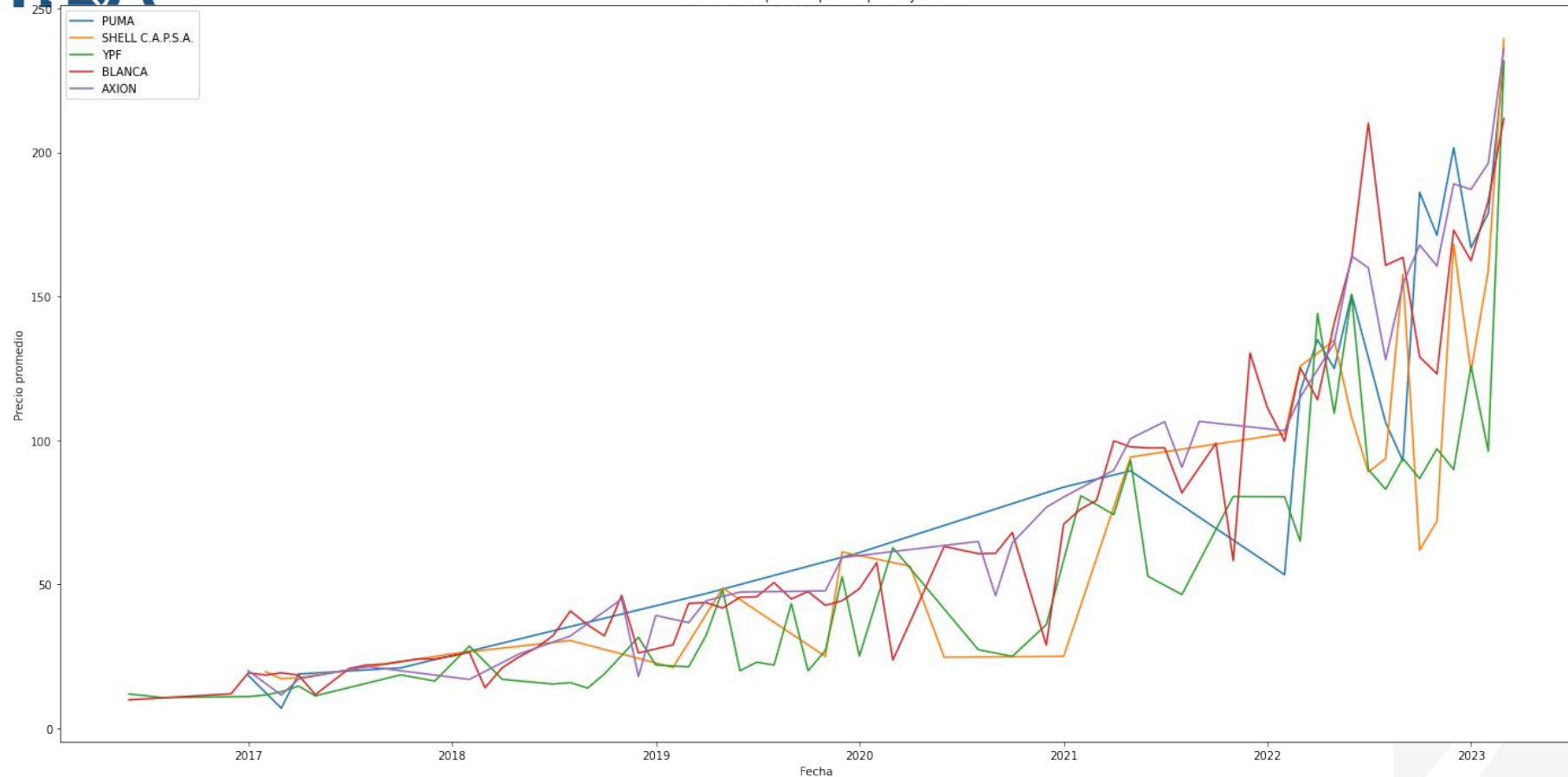


```
indice_tiempo 0
localidad      0
provincia      0
region        6389
producto       0
tipohorario    0
precio         0
fecha_vigencia 0
empresabandera 0
latitud        12
longitud       12
geojson        12
```











Instituto Tecnológico
de Buenos Aires

Conclusiones

y preguntas!