Proyecto final: pronóstico de pedidos 2026 de PyShop

Durante el curso limpiaste fechas, cambiaste de frecuencia, descompusiste la serie, mediste su estacionariedad y ajustaste Holt-Winters y SARIMA. Ahora vas a juntar todo en un pronóstico mensual de 2026 que el equipo de PyShop pueda usar para planear inventario. Se usan Python 3.12 o superior, pandas, NumPy, statsmodels 0.14 y matplotlib. El notebook `proyecto-pronostico-pedidos` contiene una solución de referencia: consúltalo al terminar, no antes.

Contexto

PyShop registra cuántos pedidos recibe cada día desde 2022. El equipo de compras necesita saber cuántos pedidos esperar cada mes de 2026, con un rango, y en particular cuánto inventario preparar para noviembre y diciembre, los meses del Buen Fin y las fiestas. Quedarse corto significa ventas perdidas; pasarse, inventario parado.

Datos

Copia estas funciones en la primera celda de tu notebook. `generar_pedidos()` crea la serie diaria sintética (con 18 días sin registro a propósito) y `cargar_co2()` la serie real de CO2 para el extra.

```
import numpy as np
import pandas as pd

SEMANA = np.array([1.18, 1.08, 1.04, 1.02, 0.98, 0.84, 0.86])
MES = np.array([0.82, 0.86, 0.95, 0.97, 1.02, 0.96, 0.95, 0.97, 0.98,


def generar_pedidos(semilla=42, con_faltantes=True):
    rng = np.random.default_rng(semilla)
    fechas = pd.date_range("2022-01-01", "2025-12-31", freq="D")
    t = np.arange(len(fechas))
    nivel = (180 + 0.09 * t) * SEMANA[fechas.dayofweek] * MES[fechas.
    buen_fin = np.zeros(len(fechas))
    for anio in range(2022, 2026):
        nov = pd.date_range(f"{anio}-11-01", f"{anio}-11-30", freq="D
        viernes = nov[nov.dayofweek == 4][2]
        dias = pd.date_range(viernes - pd.Timedelta(days=1), periods=
        buen_fin[fechas.isin(dias)] = 0.9
```

```
    nivel = nivel * (1 + buen_fin)
    ruido = rng.normal(0, 0.06, len(fechas))
    deriva = np.exp(np.cumsum(rng.normal(0, 0.003, len(fechas))))
    pedidos = np.round(nivel * deriva * (1 + ruido)).astype(float)
    s = pd.Series(pedidos, index=fechas, name="pedidos")
    s.index.name = "fecha"
    if con_faltantes:
        caidas = rng.choice(len(fechas), 18, replace=False)
        s = s.drop(s.index[caidas])
    return s


def cargar_co2():
    from statsmodels.datasets import co2
    serie = co2.load_pandas().data["co2"]
    return serie.resample("MS").mean().interpolate()


s = generar_pedidos()
print(len(s), s.index.min().date(), s.index.max().date())
# 1443 2022-01-01 2025-12-31
```

El Buen Fin de estos datos es un pico sintético de 5 días a mediados de noviembre; no representa fechas reales.

Requisitos mínimos

1. **Limpieza.** Completa la frecuencia diaria con `asfreq("D")`, reporta cuántos días faltan y cuáles, e interpola. Justifica por qué no rellenas con 0.
2. **Exploración.** Gráfica de la serie diaria con una media móvil de 30 días; medias por día de la semana, por mes y por año; total mensual con `resample("MS").sum()`. Comenta qué año creció menos.
3. **Descomposición.** `seasonal_decompose` mensual (periodo 12) aditiva y multiplicativa; elige una y justifica con cifras. Agrega un STL diario con periodo 7 y describe el patrón semanal.
4. **Estacionariedad.** ADF sobre la serie mensual y sobre sus diferencias (`diff(1)`, `diff(12)`, ambas). Grafica la ACF y señala los rezagos importantes.
5. **Partición y piso.** Entrena con 2022-2024 y prueba con 2025. Calcula MAE y MAPE de tres ingenuos: último valor, media histórica y naive estacional.
6. **Suavizamiento exponencial.** Ajusta SES, Holt y Holt-Winters aditivo y multiplicativo; reporta sus parámetros y su error en 2025.

7. **SARIMA.** Rejilla con $p, q, P, Q \in \{0, 1\}$, $d = D = 1$, $s = 12$; elige por AIC. Reporta MAE y MAPE en 2025, cuántos meses reales caen en el intervalo del 95 % y el p-valor de Ljung-Box de los residuos.
8. **Origen móvil.** Al menos tres cortes con horizonte de 6 meses; compara SARIMA, Holt-Winters y el naive estacional en una tabla.
9. **Pronóstico 2026.** Reentrena el modelo elegido con 2022-2025, pronostica los 12 meses de 2026 con intervalo del 95 % y grafícalo junto a la historia. Reporta el total anual y su crecimiento contra 2025.
10. **Recomendación de inventario.** Para noviembre y diciembre de 2026, propón un número de pedidos a cubrir usando el intervalo, no solo el pronóstico central, y explica el criterio.

Pistas de implementación

```
from statsmodels.tsa.statespace.sarimax import SARIMAX

final = SARIMAX(m, order=orden, seasonal_order=orden_est).fit(disp=Fa
f26 = final.get_forecast(12)
tabla = f26.conf_int(alpha=0.05).assign(pronostico=f26.predicted_mear
```

Referencia: el naive estacional da 15.5 % de MAPE en 2025 y Holt-Winters multiplicativo 3.4 %.

Extras opcionales

- **Repetir con CO2.** Con `cargar_co2()`, prueba con 1999-2001, compara Holt-Winters, SARIMA y naive estacional por MAE y pronostica 2002.
- **Pronóstico diario de noviembre.** Reparte el total mensual de noviembre de 2026 entre sus días usando el patrón semanal y el pico del Buen Fin de años anteriores.
- **Tendencia amortiguada.** Prueba `damped_trend=True` en Holt-Winters.

Entregables

1. **`proyecto.ipynb`**, ejecutado de principio a fin sin errores, con las gráficas a la vista.
2. **`INFORME.md`** (una página): modelo elegido y por qué, tabla de errores contra los ingenuos, resultado del origen móvil, pronóstico 2026 con intervalo, la recomendación de inventario y dos limitaciones. Los datos son sintéticos: describe lo que ves sin inventarle causas.

Rúbrica (100 puntos, más 10 de extras)

- **Limpieza y exploración (15 puntos):** faltantes detectados e interpolados, `resample` con la agregación correcta, gráficas comentadas.
- **Descomposición y estacionariedad (15 puntos):** elección aditiva/multiplicativa con cifras, STL, ADF y ACF interpretados con cautela.
- **Partición e ingenuos (10 puntos):** corte por fecha, tres ingenuos, MAE y MAPE bien calculados.
- **Modelos (20 puntos):** suavizamiento exponencial y SARIMA ajustados, parámetros reportados, selección por AIC sin mirar la prueba, residuos revisados.
- **Origen móvil (10 puntos):** tres cortes o más, tabla comparativa y conclusión.
- **Pronóstico 2026 (15 puntos):** intervalo del 95 %, gráfica, total anual y crecimiento.
- **Recomendación e informe (15 puntos):** cifra de inventario justificada con el intervalo, conclusiones con números y limitaciones honestas.
- **Extras (hasta 10 puntos):** CO2, reparto diario de noviembre o tendencia amortiguada bien interpretados.