# Proyecto: Generación de Datos Sintéticos con TimeGAN para Construcción de Portafolios Financieros

Contenido

[Proyecto: Generación de Datos Sintéticos con TimeGAN para Construcción de Portafolios Financieros 1](#_Toc196492760)

[Objetivo General 2](#_Toc196492761)

[Metodología 2](#_Toc196492762)

[Resultados Esperados 3](#_Toc196492763)

[Herramientas Usadas 3](#_Toc196492764)

[Conclusiones 3](#_Toc196492765)

Este proyecto realiza:  
- Análisis de datos históricos 2023-2024 de AAPL, MSFT, NVDA, META, AMZN, GOOG.  
- Entrenamiento de un modelo TimeGAN.  
- Generación de datos sintéticos.  
- Evaluación comparativa real vs sintético.  
- Construcción de portafolios.

## Objetivo General

Desarrollar un flujo de trabajo que permita aumentar series de tiempo financieras mediante GANs y evaluar su impacto en estrategias de portafolio.

## Metodología

1. Adquisición de Datos Reales  
- Descarga de precios ajustados diarios de las acciones seleccionadas (2023-2024).  
  
2. Cálculo de Log-Retornos  
- Transformación de precios en log-retornos diarios.  
  
3. Análisis Estadístico  
- Media, varianza, asimetría, curtosis, correlaciones.  
  
4. Visualización  
- Series de precios, retornos, histogramas y heatmaps.  
  
5. Preprocesamiento  
- Normalización de retornos (MinMaxScaler).  
- Creación de ventanas de secuencias (50 días).  
  
6. Entrenamiento de TimeGAN  
- Parámetros: secuencia de 50 días, latent\_dim=24, hidden\_dim=32, epochs=5000.  
  
7. Generación de Datos Sintéticos  
- Uso del modelo entrenado para generar secuencias nuevas.  
  
8. Inversión de Escalado  
- Recuperación de escala original de retornos.  
  
9. Reconstrucción de Precios Sintéticos  
- Aplicación acumulativa de log-retornos.  
  
10. Evaluación Comparativa  
- Histogramas, matrices de correlación, estadísticas descriptivas, distancia de Wasserstein.  
  
11. Simulación de Portafolios  
- Portafolio equi-ponderado: rentabilidad acumulada, Sharpe Ratio, máximo drawdown.

## Resultados Esperados

- Incremento del conjunto de datos.  
- Mejor modelado de riesgos mediante simulaciones plausibles.  
- Evaluación robusta de estrategias de inversión en escenarios extendidos.

## Herramientas Usadas

Python (pandas, numpy, matplotlib, seaborn, scikit-learn, tensorflow, timeseries\_gans).

## Conclusiones

El modelo TimeGAN permite generar datos sintéticos que preservan propiedades estadísticas de series reales, facilitando el stress testing de portafolios y exploración de escenarios futuros.