

# Examen final

## Python for Finance

Universidad Francisco Marroquín

19 de noviembre de 2025

Nombre y apellidos: \_\_\_\_\_

Duración: 1 hora y 30 minutos

Puntuación total: 10 puntos

Recursos permitidos: Ninguno (sin ordenador, sin apuntes)

---

### Ejercicio 1 – Lectura de código Python (2 puntos)

Para cada uno de los siguientes fragmentos de código, **escribe cuál será el resultado** de su ejecución. No es necesario justificar, solo indica el output.

1. (0.25 puntos)

```
precios = [100, 102, 98, 105, 103]
rendimientos = [precios[i] - precios[i-1] for i in range(1, len(precios))]
print(rendimientos)
```

Resultado: \_\_\_\_\_

2. (0.25 puntos)

```
import numpy as np
x = np.array([1, 2, 3, 4, 5])
y = 2 * x + 10
print(y[2])
```

Resultado: \_\_\_\_\_

3. (0.5 puntos) Explica la diferencia entre un **DataFrame** de pandas y una **Series** de pandas. ¿Cuándo usarías uno u otro?

4. (0.5 puntos)

```
import pandas as pd
df = pd.DataFrame({'precio': [100, 105, 103, 108]})
df['rendimiento'] = df['precio'].pct_change()
print(df['rendimiento'].iloc[2])
```

Resultado: \_\_\_\_\_

5. (0.5 puntos)

```
import pandas as pd
serie = pd.Series([10, 20, 30, 40, 50])
resultado = serie[serie > 25]
print(len(resultado))
```

Resultado: \_\_\_\_\_

---

### Ejercicio 2 – Pandas y manipulación de datos (2 puntos)

1. (0.5 puntos) Interpreta el siguiente código de pandas y explica qué hace cada parámetro:

```
df = pd.read_csv('data/gdp-fmt.csv', sep=";", decimal=".",
                parse_dates=["observation_date"],
                index_col="observation_date")
```

Explica específicamente qué hacen los parámetros: `sep`, `decimal`, `parse_dates` e `index_col`.

2. (0.5 puntos) Explica qué hace el método `.diff()` en pandas y qué hace el método `.pct_change()`. ¿En qué se diferencian? Proporciona un ejemplo conceptual de cuándo usarías cada uno.

3. **(0.5 puntos)** Tienes datos anuales de PIB pero necesitas obtener datos trimestrales para tu análisis. ¿Qué técnica utilizarías para generar valores intermedios? Describe conceptualmente el proceso.
4. **(0.5 puntos)** Explica qué hace el método `.describe()` en un DataFrame de pandas y qué información proporciona. Menciona al menos 3 estadísticas que incluye en su resultado.

---

### Ejercicio 3 – Regresión lineal (2 puntos)

Has construido un modelo de regresión lineal usando `statsmodels` para predecir el rendimiento de una acción basándote en el rendimiento del mercado. El siguiente es el resumen (`summary`) del modelo:

OLS Regression Results						
=====						
Dep. Variable:	rendimiento		R-squared:	0.682		
Model:	OLS		Adj. R-squared:	0.678		
Method:	Least Squares		F-statistic:	168.4		
Date:	Mon, 18 Nov 2025		Prob (F-statistic):	3.21e-23		
Time:	14:30:00		Log-Likelihood:	245.67		
No. Observations:	80		AIC:	-487.3		
Df Residuals:	78		BIC:	-482.5		
Df Model:	1					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
const	0.0015	0.002	0.750	0.455	-0.002	0.005
mercado	1.2450	0.096	12.977	0.000	1.054	1.436
-----						

1. **(0.5 puntos)** Identifica y explica el significado de `coef` de la variable `mercado` y de `const`. ¿Qué indican estos valores en el contexto del modelo? Interpreta el resultado del coeficiente de `mercado`.
2. **(0.5 puntos)** ¿Qué representa el **valor p** en la tabla? Según los resultados, ¿cuál de las dos variables (`const` o `mercado`) es estadísticamente significativa y por qué?
3. **(0.5 puntos)** Identifica el **R-squared** en la tabla. ¿Qué significa que sea 0.682? ¿Es un buen ajuste para este modelo? Justifica tu respuesta.
4. **(0.5 puntos)** ¿Qué muestra la columna `[0.025 0.975]`? ¿Qué nos dice este intervalo?

---

### Ejercicio 4 – Clasificación y métricas (2 puntos)

Has desarrollado un modelo de clasificación binaria para predecir si una empresa entrará en bancarrota (1 = bancarrota, 0 = no bancarrota). Tras evaluar el modelo obtienes la siguiente matriz de confusión:

	Predicho: No (0)	Predicho: Sí (1)
Real: No (0)	850	50
Real: Sí (1)	30	70

1. **(0.5 puntos)** Define **Precision** (precisión) en el contexto de clasificación binaria. Calcula la precisión de este modelo mostrando la fórmula y el resultado.
2. **(0.5 puntos)** Define **Recall** (sensibilidad o exhaustividad) en el contexto de clasificación binaria. Calcula el recall de este modelo mostrando la fórmula y el resultado.
3. **(0.5 puntos)** Define **F1-score** y explica qué representa. ¿Por qué es útil tener una métrica que combine precisión y recall? Calcula el F1-score para este modelo mostrando la fórmula y el resultado.
4. **(0.5 puntos)** Menciona al menos **tres modelos de clasificación** diferentes que podrías utilizar para este problema. Ventajas y desventajas de cada uno.

**Extra:** ¿Qué otras métricas aparte de *precision*, *recall* y F1 se podrían usar? Sólo se pide nombrarlas, no explicarlas.

---

### Ejercicio 5 – Series temporales con Darts (2 puntos)

Estás analizando una serie temporal de ventas mensuales de una empresa utilizando la librería **darts**.

1. **(0.5 puntos)** Explica por qué en series temporales es fundamental hacer un **split temporal** (entrenamiento con datos pasados, test con datos futuros) en lugar de un split aleatorio como se hace en otros problemas de machine learning.
2. **(0.5 puntos)** Menciona tres tipos de modelos que podrías usar con **darts** para hacer predicciones. Explica cada uno brevemente.
3. **(0.5 puntos)** Explica qué miden las siguientes métricas de error en series temporales:
  - MAE
  - RMSE
  - MAPE

¿Cuál de estas métricas sería más apropiada si quieres que los errores se expresen en términos porcentuales?

4. **(0.5 puntos)** Si una serie temporal muestra tendencia y estacionalidad, ¿es buena idea usar un modelo **NaiveMean**? ¿Cuál o cuáles propondrías en su lugar? Justifica ambas.

**Extra:** Responde brevemente: ¿y si haces preprocesado, sería buena idea usar **NaiveMean**? ¿Qué preprocesado harías?