

# BUSINESS INTELLIGENCE PER I SERVIZI FINANZIARI

## PROGETTO DI LABORATORIO IN PYTHON

Versione del 3 aprile 2025

[antonio.candelieri@unimib.it](mailto:antonio.candelieri@unimib.it)

[silvio.bencini@unimib.it](mailto:silvio.bencini@unimib.it)

### 1. Informazioni generali

Il progetto è finalizzato a mostrare le capacità acquisite in:

- acquisizione dei dati;
- visualizzazione;
- analisi esplorativa delle serie;
- analisi esplorativa di portafoglio;
- previsione di rendimenti,

utilizzando le librerie di Python introdotte nei laboratori (Pandas, Matplotlib, Numpy, ScikitLearn, StatsModels eccetera) e le funzioni principali di ciascuno.

Gli studenti sono liberi di utilizzare altre librerie di Python (purché compatibili con versioni di Python uguali o superiori a 3.8).

Il progetto è individuale.

### 2. Dati da utilizzare

Lo studente può scegliere fra due insiemi di dati:

- almeno 6 azioni del mercato americano facenti parte dell'indice S&P 500 da analizzare, rappresentative di 3 diversi settori (per esempio banche, automobili, minerario, telecomunicazioni ...). La scelta degli emittenti deve essere basata su un'analisi di notizie finanziarie tratte dai principali siti web;
- almeno 6 ETF settoriali facenti parte dell'indice S&P500 o simili (MSCI USA).

Il periodo da utilizzare è 31.5.2014 – 31.5.2024

### 3. Documenti di progetto

Lo studente deve produrre:

- un Jupyter Notebook contenente il codice utilizzato per il lavoro, adeguatamente commentato;
- una presentazione in ppt di 10/15 minuti per l'illustrazione dei principali risultati;
- una relazione completa divisa nei seguenti capitoli (vedi schema di dettaglio allegato):
  - o sommario dei dati utilizzati;
  - o statistiche descrittive;
  - o analisi di previsione;
  - o beta di ciascun titolo;
  - o creazione di un portafoglio o simulazione di strategia dinamica
  - o conclusioni.

I documenti di progetto devono essere inviati una settimana prima della data dell'esame a [silvio.bencini@unimib.it](mailto:silvio.bencini@unimib.it) e [antonio.candelieri@unimib.it](mailto:antonio.candelieri@unimib.it)

#### 4. Schema di documento finale

##### 1. Sommario dei dati utilizzati

- a. Breve descrizione di ciascun titolo o ETF selezionato e motivazione della scelta (per esempio notizie rilevanti, trend osservati eccetera);
- b. Funzioni utilizzate per scaricare i dati da Yahoo! Finance o da altri siti (es. Fama-French).
- c. Funzioni utilizzate per la fusione (se necessario) delle serie in un unico DataFrame;
- d. Presentazione dei dati con un grafico e le prime righe del DataFrame.

##### 2. Statistiche descrittive

- a. Calcolare il rendimento cumulato e composto annuo di ciascun titolo nel periodo;
- b. Calcolare i rendimenti semplici e logaritmici e visualizzarli in un grafico;
- c. Commentare:
  - i. che cosa hanno in comune le serie storiche?
  - ii. c'è una correlazione positiva fra società dello stesso settore
  - iii. ci sono momenti di rendimenti molto lontani dalla media? Se sì cercate le notizie che potrebbero spiegarli.
- d. Creare grafici diagnostici a 3 sezioni (istogramma e kernel density, boxplot, qq-plot) per ciascuna serie di rendimenti e commentare (si notano differenze nella dispersione dei rendimenti, i rendimenti sono distribuiti normalmente? Ci sono outliers?)
- e. Calcolare statistiche descrittive univariate annualizzate (media, varianza, deviazione standard, asimmetria, curtosi) per ogni serie di rendimenti e commentare.
  - i. Quali azioni (ETF) hanno il rendimento più basso e più alto?
  - ii. quali azioni (ETF) hanno la deviazione standard più alta o più bassa?
  - iii. come si evolvono nel tempo rendimento e volatilità?
  - iv. quale azione(ETF) ha la distribuzione di rendimenti più vicina o lontana dalla normale?
- f. Calcolare la matrice di varianze/covarianze e di correlazione dei rendimenti mensili. Sulla base della matrice di correlazioni:
  - i. Quali sono i titoli più correlati?
  - ii. Quali i meno correlati?
- g. Fare il grafico dell'andamento nel tempo delle correlazioni fra i titoli e i grafici di dispersione (scatter plots) dei rendimenti dei due titoli all'interno dello stesso settore o di ciascun ETF con l'indice di mercato, dai quali risulti la maggiore o minore la correlazione storica dei rendimenti
  - i. Commentare le relazioni e il loro andamento nel tempo;
  - ii. Come cambia la correlazione fra le azioni o fra ETF e mercato nel tempo?
  - iii. Come cambiano le correlazioni in funzione dei rendimenti?
  - iv. La dispersione dei punti negli scatter plot conferma o no la relazione lineare fra i due rendimenti?

##### 3. Analisi di previsione

- a. Costruire un modello di previsione ARIMA) per prevedere l'andamento di ciascun strumento finanziario, usando:
  - i. n (80) mesi come training set
  - ii. m (30) mesi come test set

- iii. I (10) mesi per la previsione, che può essere svolto sia sull'intero periodo sia ricorsivamente mese dopo mese
  - iv. Utilizzare gli ultimi (10) mesi per confrontare le previsioni con i valori effettivi
  - v. Verificare l'errore medio della previsione
- 4. Strategie di trading e backtesting – Analisi tecnica
  - a. Costruire una strategia di trading basata su un algoritmo a scelta che segnali l'acquisto o la vendita di un titolo o indice di borsa e farne il backtesting
  - b. Misurare l'efficienza della strategia rispetto alla detenzione del titolo per tutto il periodo ("Buy & Hold");
  - c. Utilizzare variabili di mercato ma non collegate ai prezzi passati del titolo (volume, VIX, andamento dell'indice o variabili non di mercato (Google Trends) - **Opzionale**
- 5. CAPM
  - a. Calcolare il beta di ciascun titolo o ETF rispetto al mercato (indice S&P 500, ticker Yahoo Finance ^GSPC)
  - b. Utilizzare il beta per calcolare il rendimento atteso annuo del titolo (utilizzare una propria previsione o stima del rendimento dell'indice S&P500 e rendimento di mercato del T-Bill a 1 anno).
  - c. Calcolare l'esposizione di ciascun titolo o ETF ai fattori di rischio Fama-French
- 6. Strategie di trading e backtesting – Strategie dinamiche
  - a. Simulare almeno una strategia di investimento sui titoli in portafoglio fra le seguenti: Equal weight, Stop Loss, CPPI, TPPI
  - b. Confrontare la strategia dinamica con quella "Buy & Hold"
  - c. Discutere i parametri che determinano il risultato della strategia e valutarne possibili modifiche
- 7. Costruzione di portafoglio
  - a. Costruire il portafoglio ottimale in termini di media-varianza utilizzando i primi 108 mesi di dati, sia con metodo analitico sia con metodo di simulazione, utilizzando sia i rendimenti passati sia i rendimenti attesi costruiti nella parte 6
  - b. Calcolare il beta del portafoglio rispetto al mercato
  - c. Confrontare il rendimento del portafoglio ottimale con quello effettivo. Per "portafoglio effettivo" si intende un portafoglio composto dai sei titoli oggetto di analisi con peso uguale fra di loro.

I punti 6 e 7 sono alternativi.