BUSINESS INTELLIGENCE PER I SERVIZI FINANZIARI

PROGETTO DI LABORATORIO IN PYTHON

antonio.candelieri@unimib.it

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 superiori a 3.8).

Il progetto è individuale.

2. Dati da utilizzare

Lo studente deve scegliere almeno 6 azioni del mercato americano 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.

Il periodo da utilizzare è 30.11.2011 – 30.11.2021.

3. Documenti di progetto

Lo studente deve produrre:

- un Jupytes 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 conclusioni.

I documenti di progetto devono essere inviati una settimana prima della data dell’esame a silvio.bencini@unimib.it.

1

BUSINESS INTELLIGENCE PER I SERVIZI FINANZIARI

PROGETTO DI LABORATORIO IN PYTHON

Schema di documento finale

1. Sommario dei dati utilizzati

a. Breve descrizione di ciascun titolo 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 i rendimenti semplici e composti e visualizzarli in un grafico;

b. 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.

c. Presentare i rendimenti con istogrammi e confrontare la dispersione dei rendimenti dei diversi titoli

d. Creare grafici diagnostici a 4 sezioni (istogramma, kernel density, boxplot, qq-plot) per ciascuna serie di rendimenti e commentare (i rendimenti sono distribuiti normalmente? Ci sono outliers?)

e. Calcolare statistiche descrittive univariate (media, varianza, deviazione standard, asimmetria, curtosi) per ogni serie di rendimenti e commentare.

i. Quali azioni hanno il rendimento più basso e più alto?

ii. quali azioni hanno la deviazione standard più alta o più bassa?

iii. come si evolvono nel tempo rendimento e volatilità?

iv. quale azione ha la distribuzione di rendimenti più vicina o lontana dalla

normale?

f. Calcolare la matrice di varianze/covarianze dei rendimenti e commentare le relazioni fra i diversi titoli. Qual é più rischioso?

g. Calcolare la matrice di correlazione dei rendimenti.

i. Quali sono i titoli più correlati?

ii. Quali i meno correlati?

h. Fare il grafico dell’andamento nel tempo delle correlazioni fra i titoli e i grafici di dispersione (scatter plots) delle correlazioni medie.

i. Commentare le relazioni e il loro andamento nel tempo;

ii. Come cambia la correlazione fra le azioni 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, SVM o altro) per prevedere i prezzi o rendimenti di ciascun strumento finanziario, usando:

i. n (80) mesi come training set

ii. m (30) mesi come test set

iii. l (10) mesi per la validazione

2

iv. Utilizzare gli ultimi l (10) mesi per confrontare le previsioni con i valori effettivi 4. Stategie di trading e backtesting

a. Costruire una strategia di trading basata su un algoritmo a scelta che segnali l’acquisto o la vendita di un titiolo o indice di borsa e farne il backtesting

b. 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 rispetto al mercato (indice S&P 500, ticker Yahoo Finance ^GSPC)

b. Calcolare l’esposizione di ciascun titolo ai fattori di rischio Fama-French c. Utilizzare il beta per calcolare il rendimento atteso

6. 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 3

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.

3