

Obiettivo del progetto:

Il progetto mira ad analizzare le serie storiche dei prezzi delle commodities fornite dal provider World Bank tramite la piattaforma DBNomics. Attraverso l'API di DBNomics, verrà estratto un dataset contenente le serie d'interesse, che saranno successivamente analizzate.

Requisiti:

Per l'elaborazione del progetto è necessario installare nel proprio ambiente di lavoro i seguenti pacchetti

Procedura di estrazione del dataset:

Per estrarre il dataset utilizzando l'API, sarà necessario installare il client messo a disposizione da DBNomics, che può essere ottenuto tramite il pacchetto Python chiamato "dbnomics". Una volta installato, sarà necessario configurare l'URL corretto utilizzando la guida fornita dallo Swagger di DBNomics, accessibile all'indirizzo: <https://api.db.nomics.world/v22/apidocs>

Pulizia e trasformazione dei dati:

Dopo aver estratto il dataset originale (DF1), sarà eseguito un lavoro di pulizia e trasformazione per migliorare la sua leggibilità. Saranno aggiunte nuove colonne al dataframe originale, ottenute dalla scomposizione della colonna "description", che contiene il nome, l'indicizzazione dei prezzi reali/nominali e l'unità di misura.

Questa operazione genererà un nuovo dataframe (DF2) con una struttura più chiara e consentirà la selezione delle serie da analizzare.

Filtro delle serie:

Per facilitare la scelta delle serie da analizzare, sarà implementato un filtro basato su parola chiave.

Ci saranno due domande poste all'operatore:

- 1) Scegli il numero di serie da analizzare.
- 2) Inserisci il numero corrispondente alla serie da analizzare.

Le risposte a queste domande forniranno l'elenco dei codici delle serie selezionate per l'analisi.

Chiamata API per ottenere le osservazioni:

Successivamente, verrà effettuata una nuova chiamata API al provider utilizzando i codici delle serie selezionate, al fine di ottenere un nuovo dataset contenente le osservazioni corrispondenti.

Analisi dei dati:

Una volta ottenuto il dataframe con le osservazioni, saranno estratti gli indicatori statistici principali, come la media, i valori minimi e massimi, la distribuzione dei quantili e la covarianza delle serie. Queste analisi forniranno informazioni sul comportamento e sulle relazioni tra le diverse serie.

Salvataggio dei risultati:

I dataframe ottenuti e i risultati delle analisi saranno salvati nella cartella "OUT". Inoltre, sarà possibile scegliere un percorso di salvataggio alternativo, se necessario.

Rappresentazione grafica:

I dati elaborati saranno utilizzati per generare rappresentazioni grafiche delle serie, utilizzando grafici a linee. Inoltre, verranno creati grafici a barre per visualizzare la media e la varianza delle serie. Infine, verrà generata una rappresentazione grafica della covarianza tra le serie per analizzare le relazioni tra di esse.

Struttura dei file:

Il progetto è composto da un jupyter notebook in python nella cartella principale, dove verranno eseguite le analisi e le elaborazioni dei dati.

All'interno della stessa cartella principale, sarà presente un'altra cartella

Modulo_Agostino_Fontana, contenente i file "init.py" e "modulo_agostino_fontana.py". Il file "modulo_agostino_fontana.py" contiene le funzioni personalizzate sviluppate dall'autore, che verranno utilizzate nel notebook principale.

Tutti gli output generati, compresi i dataframe e le rappresentazioni grafiche, saranno salvati nella cartella "OUT" all'interno della cartella principale.