

Esame Laboratorio di Programmazione I

27/06/2025

Svolgimento

Chiamate il file in cui scrivere il vostro codice `esame.matricola.py` (la matricola è SMnumeri o ECnumeri). All'inizio del file scrivete un commento con il vostro nome e numero di matricola. Consegnate un codice che compila, se il codice non compila non sarà valutato!!! Commentate bene il codice descrivendo quello che fate! Hint: fate un pezzo alla volta e testate il vostro codice per vedere che funziona nel modo corretto.

Dataset Il file `GlobalLandTemperaturesByMajorCity.csv` contiene misurazioni mensili della temperatura media da 1849 al 2013, per migliaia di città raccolte. Le colonne rilevanti sono:

- `dt` (YYYY-MM-DD): la data della misurazione,
- `AverageTemperature`: la temperatura media mensile (in gradi Celsius),
- `City`: la città in cui è stata effettuata la misurazione.

Il dataset contiene mooolte righe e presenta dati mancanti e valori nulli o non affidabili.

1 (10 punti) Lettura e filtraggio dati

Scrivere la classe `CSVTimeSeriesFile`:

- (2 punti) La classe deve essere istanziata con il nome del file tramite la variabile `name`.
- (8 punti) Deve avere un metodo `get_data` che prende in input il nome della città (ad esempio Rome, nota che i nomi delle città sono in inglese), e restituisce una lista di coppie `[data, temperatura]`, dove il primo elemento è la data (sotto forma di stringa) e il secondo la temperatura media mensile (sotto forma di float), per tutti e SOLO i valori della città in input validi, quindi scritti come valori di tipo float.

Esempio d'uso:

```
time_series_file = CSVTimeSeriesFile(name="GlobalLandTemperaturesByCity.csv")
time_series_italy = time_series_file.get_data((city="Rome"))
```

2 (10 punti) Calcolo del coefficiente angolare

Definire la funzione `compute_slope(time_series, first_year, last_year)` che prende in input una serie temporale, anno iniziale e finale dell'intervallo di anni da considerare (inclusi) e ritorna il coefficiente angolare della retta di regressione sugli anni considerati. Gli anni dell'intervallo sono inseriti come tipi interi.

La funzione deve in particolare:

- (2 punti) Raggruppare le temperature per anno nel range specificato (inclusivo), considerando solo gli anni per cui sono disponibili almeno 6 misurazioni mensili valide.
- (2 punti) Calcolare la media annuale per ogni anno valido nell'intervallo dato.
- (2 punti) Calcolare la media di tutti gli anni \bar{x} , e il valore medio delle temperature medie annuali \bar{y} , cioè:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{e} \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

dove x_i sono gli anni, y_i sono le temperature medie annuali per l'anno x_i , ed n è il numero di anni validi (quindi con almeno 6 misurazioni) all'interno dell'intervallo.

- (4 punti) Calcolare e poi ritornare il coefficiente angolare m applicando la seguente formula.

$$m = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

3 (10 punti) Validazione input ed eccezioni

Le eccezioni da alzare in caso di input non corretti o casi limite devono essere istanze di una specifica classe `ExamException`, che dovete definire nel codice come segue, senza modifica alcuna (copia-incollate le due righe):

```
class ExamException(Exception):  
    pass  
  
... e che poi userete come una normale eccezione, ad esempio:  
  
    raise ExamException("Errore: impossibile aprire il file")
```

Per l'esame dovete gestire obbligatoriamente i seguenti casi:

- (2 punti) La classe `CSVTimeSeriesFile` deve controllare l'esistenza del file in `get_data` e, nel caso il file non esista, alzare un'eccezione.
- (2 punti) Se il nome della città dato in input non è presente nel file, si deve alzare una eccezione in cui viene detto `"Errore: il nome della città non è presente nel file"`.
- (2 punti) Se n risulta uguale a zero o il denominatore nel calcolo del coefficiente angolare risulta uguale a zero sollevare un'eccezione.
- (2 punti) Se l'intervallo di anni fornito è vuoto o non valido, sollevare un'eccezione.
- (2 punti) Se un anno contiene meno di 6 valori validi, bisogna ignorarlo (non si tratta di un errore, ma di un caso da gestire).