

Master DS 2022/2023

Progetto: Guarda se fa freddo

Studente: Luca Garbin

Il progetto “Guarda se fa freddo” prevede l’analisi del dataset:

“GlobalLandTemperaturesByMajorCities.csv” che contiene la temperatura registrata nelle maggiori città del mondo dal 1743 ai giorni nostri.

Lo scopo della prima parte del progetto è quello di fornire una rappresentazione grafica dei dati, dividendo le varie osservazioni in periodi storici definiti, evidenziando le maggiori escursioni termiche osservate.

La prima fase ha riguardato l’analisi del dataset prima menzionato allo scopo di cercare eventuali dati mancanti e capire se adottare delle tecniche di compensazione.

Da quest’analisi è emerso che il dataset, presentava circa il 4,6% di dati mancanti riguardanti la colonna delle temperature medie.

Vista la bassa percentuale, ho deciso di ignorare ed escludere quelle osservazioni dal computo totale.

La fase successiva, ha riguardato la definizione dei periodi storici. Visto che il dataset varia dal 1743 al 2014 (271 anni), ho deciso di prendere in considerazione un periodo di 50 anni, in modo da ottenere 6 classi in cui raggruppare le osservazioni.

Per ogni città presente e per ognuno dei 6 periodi ricavati, ho calcolato la media e la variazione della temperatura tramite la deviazione standard. Questo al fine di ottenere, per ognuno dei periodi storici, le città che hanno subito la maggiore escursione termica.

Il risultato è stato poi raggruppato in una tabella per una più immediata visualizzazione.

Per la parte di visualizzazione grafica dei risultati, ho scelto di utilizzare un barplot per ognuno dei periodi che contiene sull’asse delle ascisse le 10 città con più alta escursione termica nel periodo considerato, mentre sulle ordinate la temperatura media del periodo stesso.

Da un’attenta osservazione, si può giungere al risultato che, le città con più elevata escursione termica si trovano nel continente asiatico e più precisamente in Cina, nella regione a nord che comprende la capitale Pechino e le città a lei più prossima.

L’unica esclusione, riguarda il primo periodo storico, che va dal 1743 al 1792, in cui le città cinesi non sono presenti tra le città a maggiore escursione termica per mancanza di dati risalenti a quel periodo.

Nel grafico a linee successivo è possibile vedere la variazione della temperatura media, mese dopo mese, di alcune delle maggiori città per ogni periodo considerato.

Si può notare, ad esempio, come città asiatiche abbiano una variazione di temperatura notevole nel corso dell'anno. L'esempio più eclatante è la città cinese di Harbin, situata nel nord-est del paese, la cui temperatura media, varia da -20 gradi nel mese di gennaio ai più di 20 gradi del mese di luglio.

La seconda parte del progetto, prevede il viaggio di un individuo freddoloso tra Pechino e Los Angeles passando attraverso le città più vicine rispetto a dove si trova e che la temperatura maggiore.

Per ottenere questo risultato, ho implementato un algoritmo di tipo greedy, che prevede l'ottimizzazione di più condizioni obiettivo, che siano contemporaneamente valide rispetto alla posizione assunta di volta in volta dal nostro viaggiatore freddoloso.

Le condizioni considerate sono la minimizzazione della distanza tra la posizione corrente del viaggiatore con la città più vicina ma con temperatura maggiore. Al tempo stesso è importante minimizzare la distanza tra la futura città visitabile e la destinazione finale: questo, quando possibile, permette di convergere verso la meta finale, che è Los Angeles.

L'assunzione fatta, riguarda il concetto fondamentale, che il viaggiatore, non possa visitare la stessa città per più di una volta soltanto.

Per calcolare la distanza in chilometri tra due punti cardinali, mi sono avvalso della libreria Geopy, che mi ha permesso di ottenere la distanza reale, tenendo in considerazione la sfericità del globo terrestre. Per le temperature, si è sfruttato la media nel periodo considerato, calcolata nella prima parte del progetto.

In conclusione, è possibile notare tutto il percorso che esegue il viaggiatore, che salta da un continente all'altro fino a giungere alla destinazione.

Vengono toccate così tante città per via dei parametri stringenti, imposti dall'algoritmo.