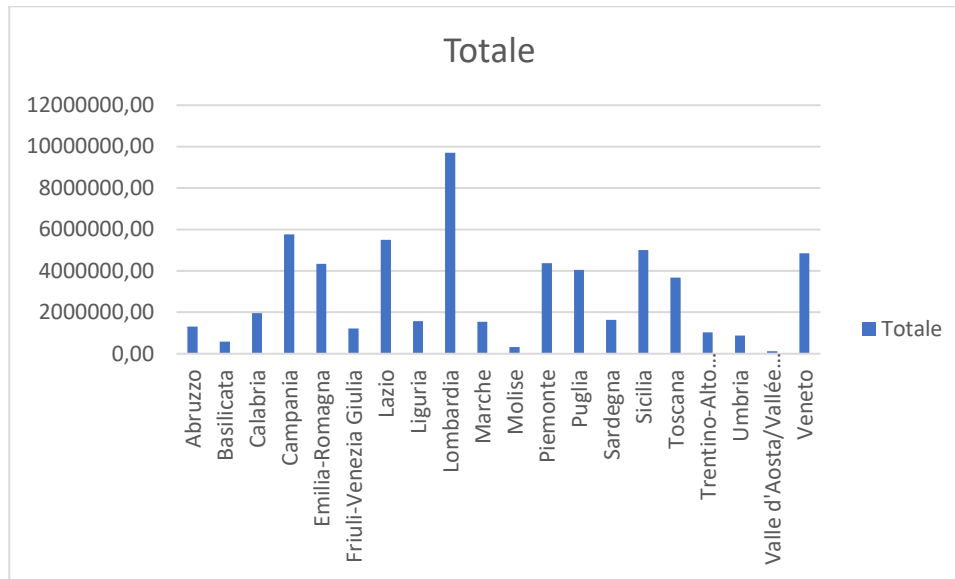


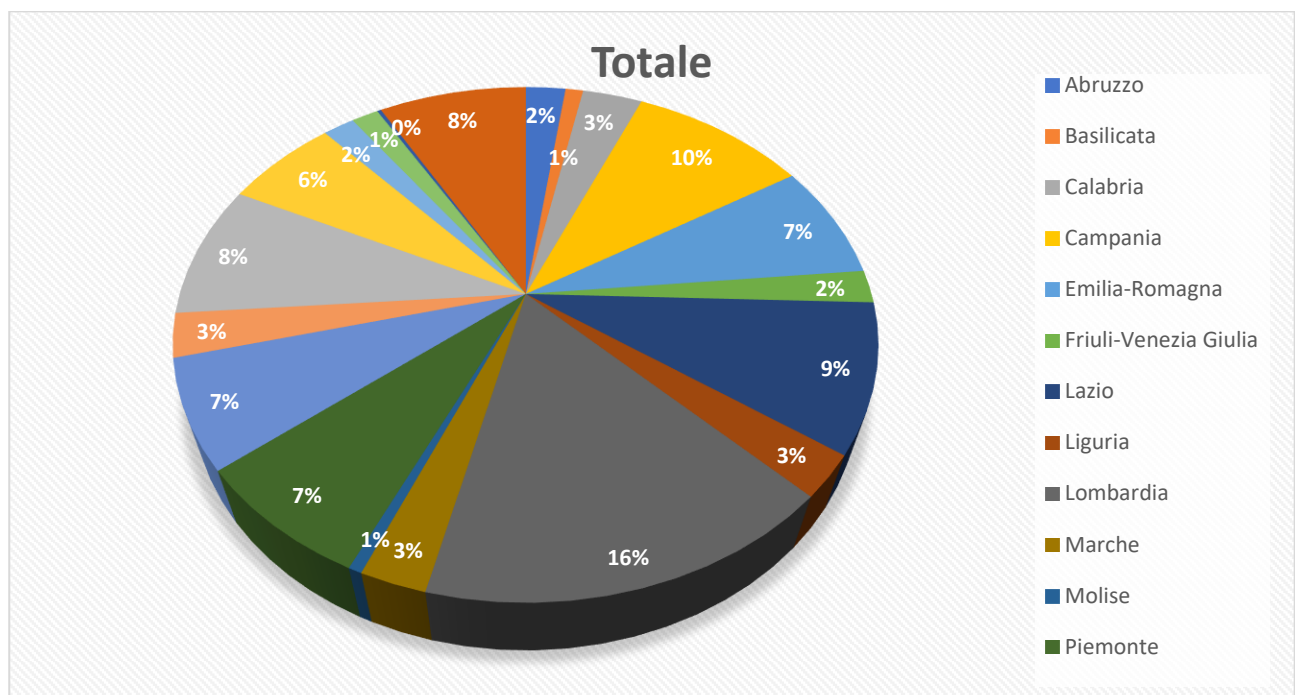
Analisi dati – Miriam Passaro

Per l'analisi dati, ho deciso di partire analizzando prima ogni singolo dataset a disposizione in modo da ricavarne i dati più utili per un confronto.

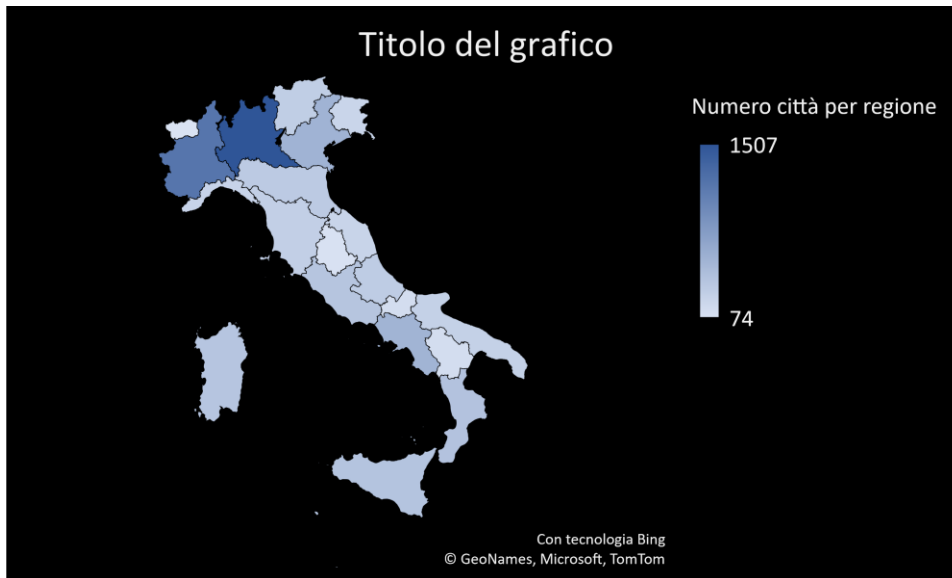
Nel dataset comuni ho deciso di fare un'analisi della distribuzione della popolazione in base alla regione e non alle singole città. Ho calcolato minimo e massimo e ho poi raggruppato la distribuzione in un grafico a colonne. La regione con il minor numero di abitanti è la Valle D'Aosta mentre quella con il maggior numero di abitanti è la Lombardia. Non avrebbe senso fare la distribuzione media dei valori poiché la distribuzione è talmente diversificata che appiattirebbe i risultati.



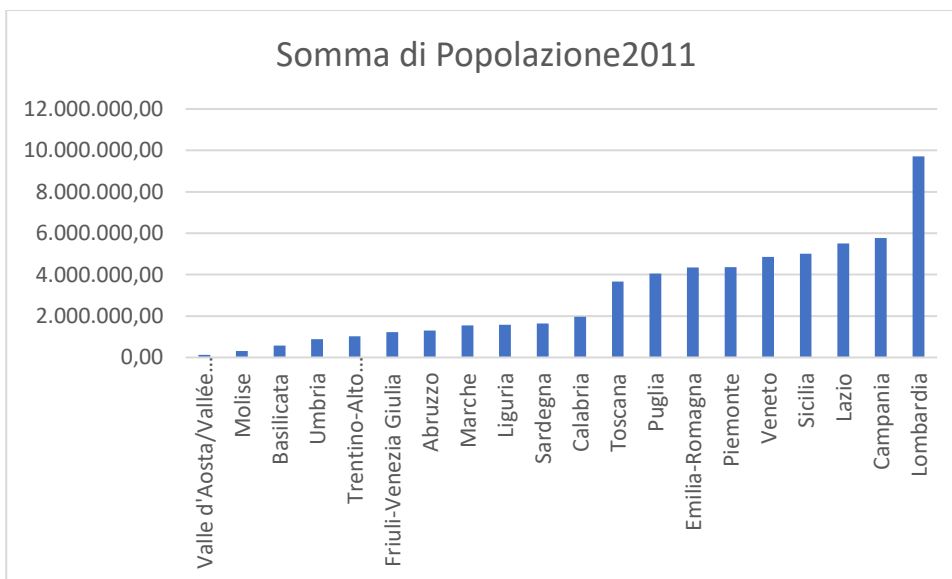
Ho anche confrontato i valori in percentuale, raggruppandoli in un grafico a torta. In questo modo è immediatamente intuibile come ogni regione contribuisce alla popolazione totale dell'Italia.

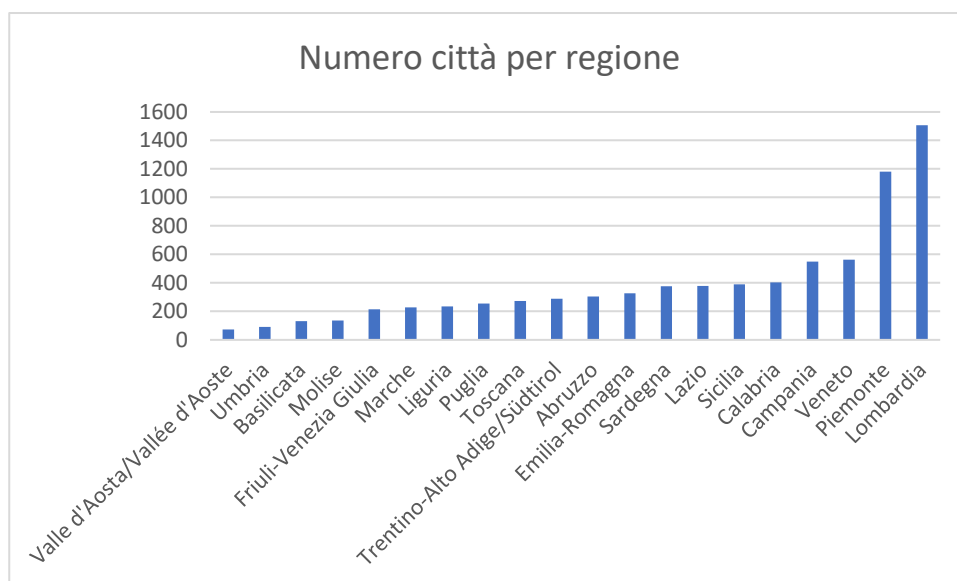


Infine, ho deciso contare il numero di città per regione presenti nel dataset. È immediatamente visibile che la Lombardia è la regione con il maggior numero di città. Con il grafico non è tuttavia immediatamente comprensibile quale regione ha meno città, tuttavia, calcolando il minimo, risulta avere il minor numero proprio la Valle D'Aosta. Dunque, c'è una corrispondenza tra la regione più popolosa e quella con maggior numero di città, nonché tra regione meno popolosa e quella con il minor numero di città.



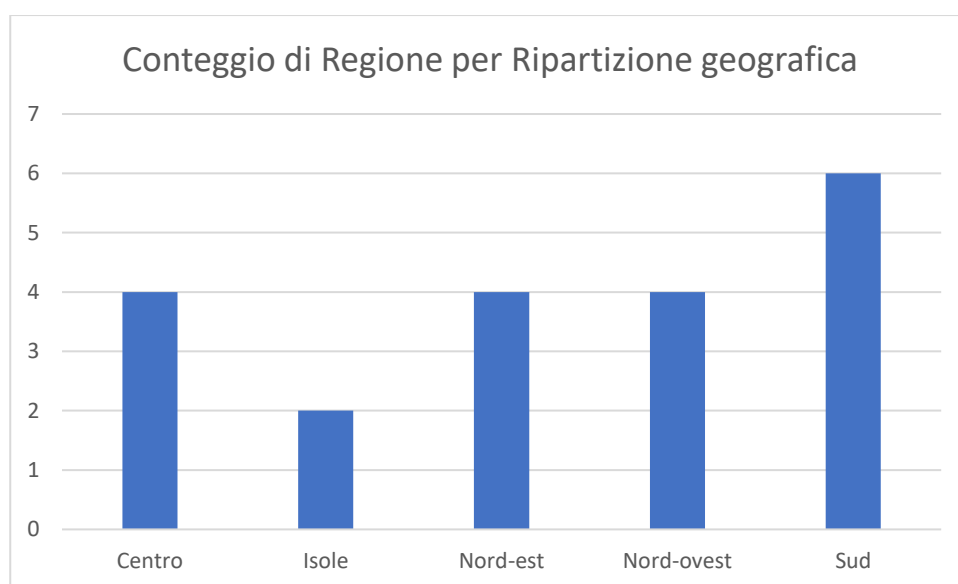
Lo stesso discorso non è valido per i valori intermedi:





Dataset della Ripartizione Geografica

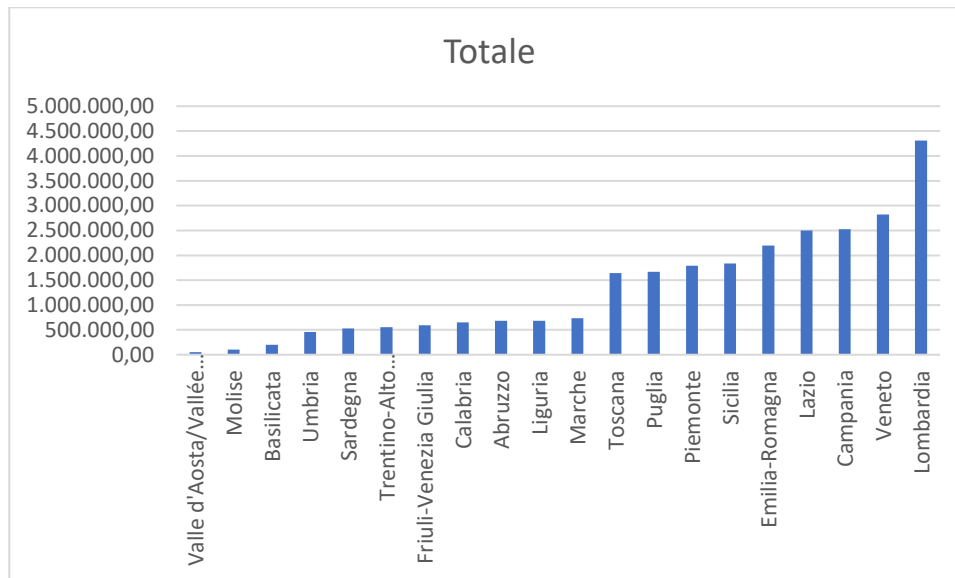
Per quanto riguarda il dataset della ripartizione geografica, mi sono limitata ad analizzare la ripartizione geografica delle regioni per aree. Ottenendo i seguenti risultati:



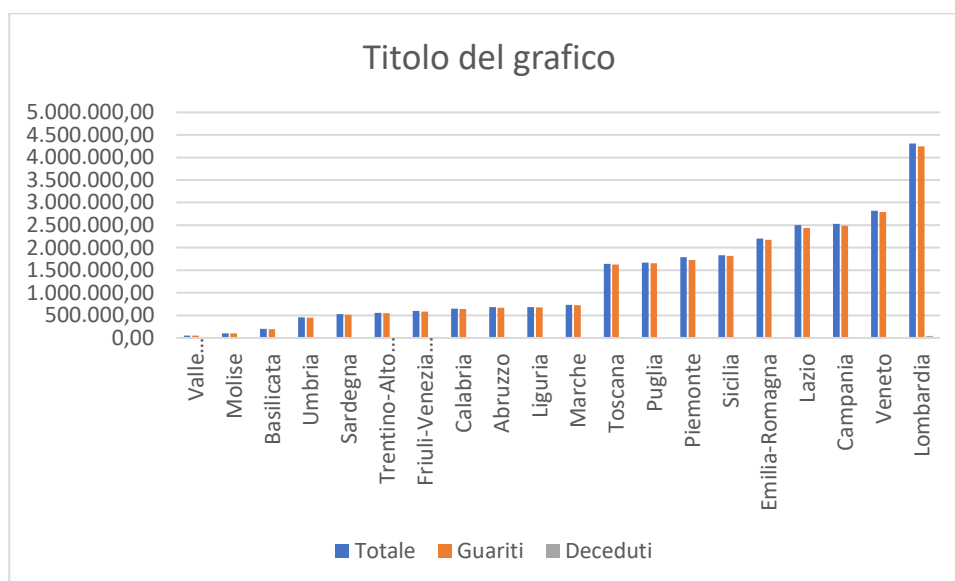
Ripartizione geografica	Conteggio di Regione
Centro	4
Isole	2
Nord-est	4
Nord-ovest	4
Sud	6

Grafico dati Covid Regioni

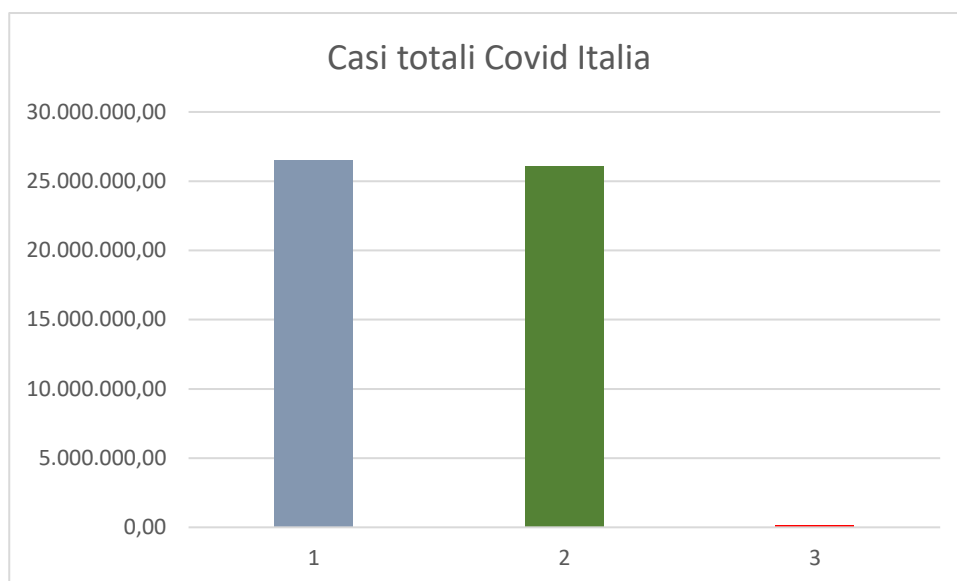
Per questo dataset ho voluto prima di tutto calcolare il tasso di decesso e guarigione per singola regione. Il tasso di decesso più alto risulta essere quello della Valle D'Aosta mentre quello con il più basso la Campania. Per quanto riguarda la distribuzione dei casi totali, risulta essere così ripartita tra le regioni:



Mentre il rapporto tra numero di casi totali per singola regione e numero di guarito e decessi risulta essere il seguente:

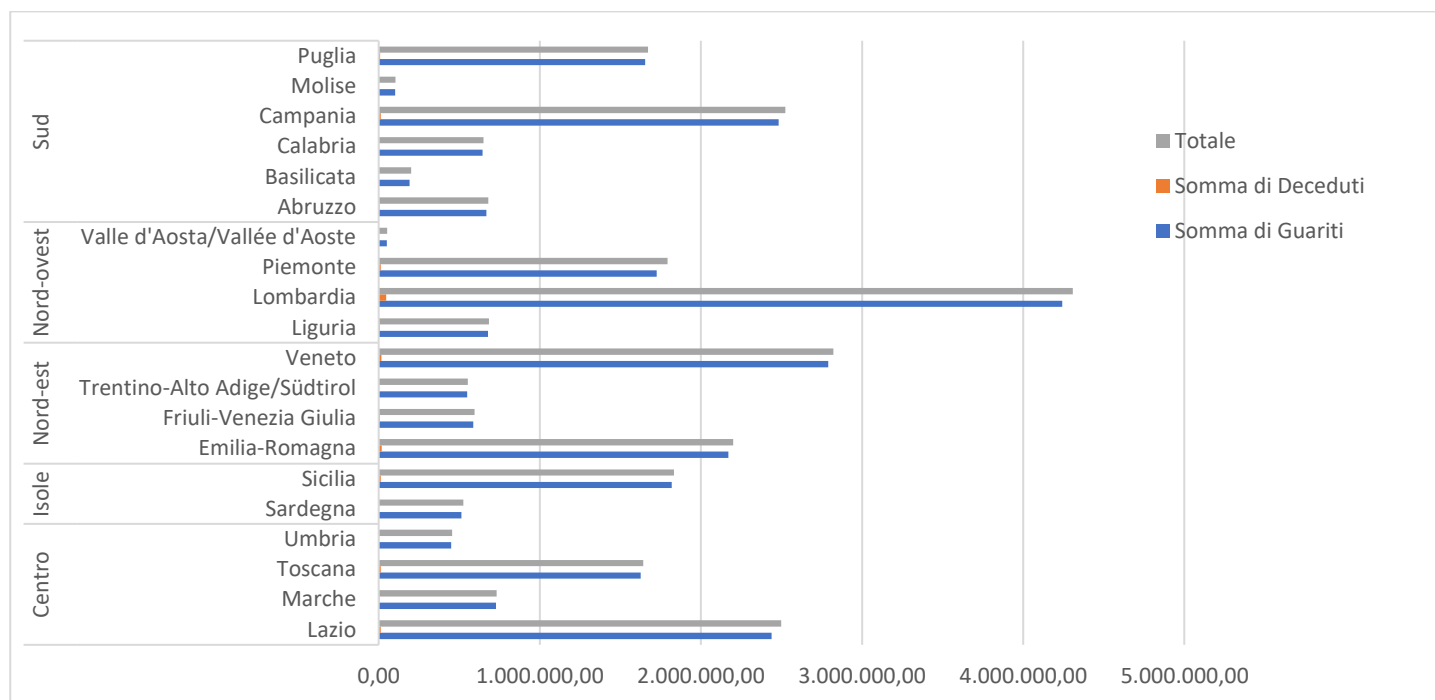


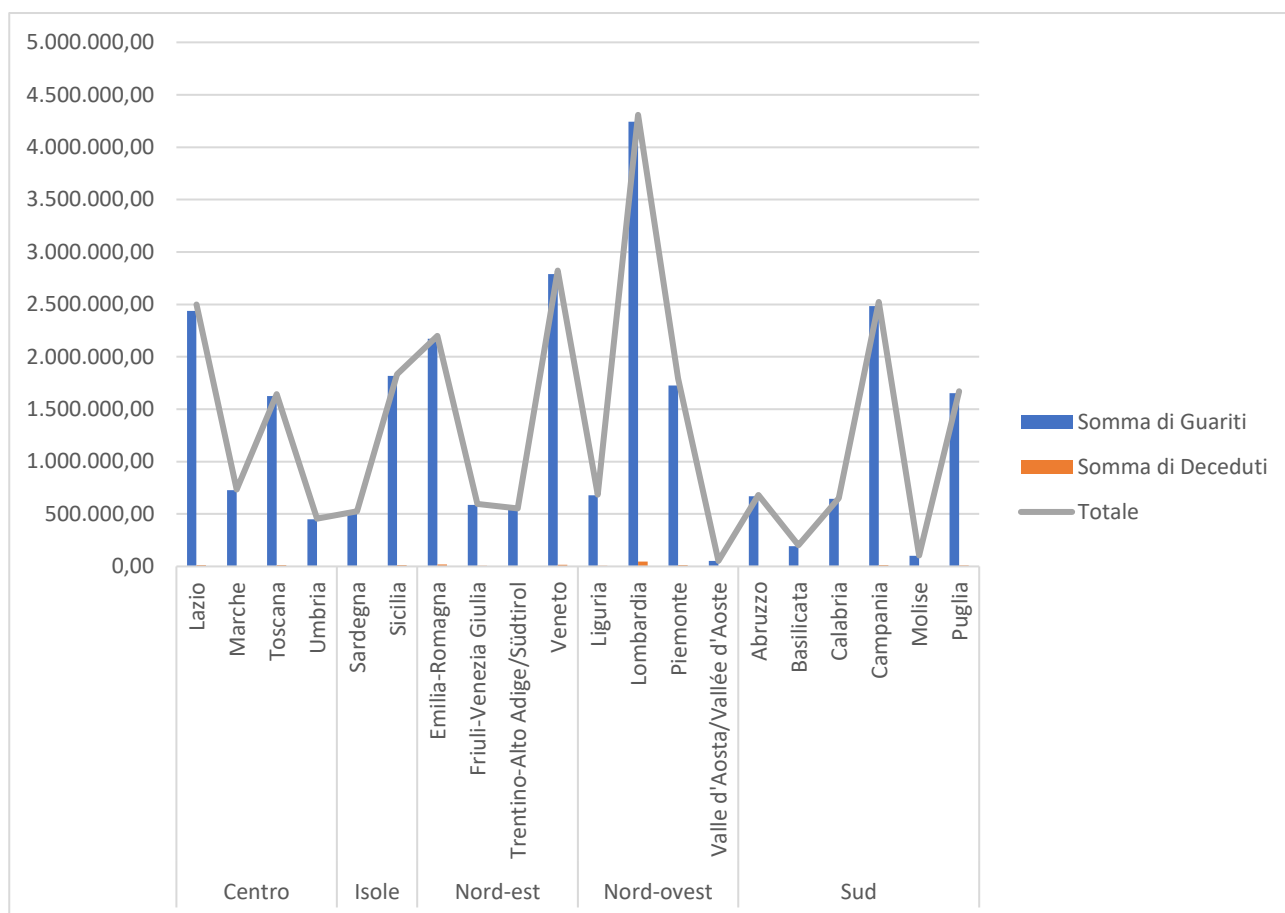
Il numero di decessi è talmente esiguo rispetto agli altri due valori che non trova una rappresentazione grafica nell'istogramma. Ho poi calcolato la somma totale dei casi in tutta Italia e l'ho messa a confronto con il numero totale di guariti e deceduti.



In blu è possibile vedere il numero totale di malati di covid, in verde il numero di guariti e in rosso il numero di deceduti.

Ho poi messo in relazione i dati del dataset sul Covid con la distribuzione delle aree geografiche (Nord-Sud e Centro) per capire come sono distribuiti i dati di casi di Covid, guariti e deceduti in relazione a delle macroaree geografiche.





L'area del nord-ovest risulta essere quella con il più alto numero di casi totali e anche di guariti e deceduti.