



UNIVERSITÀ POLITECNICA DELLE MARCHE

Corso di Project Management e supporto alle decisioni

LUCA MARCIANESI
YASSIR FLAVIO SUAREZ SANCHEZ
JUAN MANUEL CAMPOS SOLANO
RICCARDO ANGELINI

Indice

1	Introduzione	2
1.1	Data Analysis Life Expectancy Dataset	2
1.2	Sleep Health and Lifestyle Dataset	4
2	Qlik	5
2.1	Introduzione	5
2.2	Caricamento dei dati	5
2.3	Data Analysis: Life Expectancy (WHO) Fixed	6
3	Tableau	21
3.1	Introduzione	21
3.2	Gestione dei dati	21
3.3	Rappresentazioni dei dati	22
3.4	Data Analysis Life Expectancy Dataset	22
3.5	Analisi sulla copertura del vaccino contro il morbillo e correlazioni con la mortalità nei bambini	23
3.6	Analisi sul tasso di mortalità nei bambini con meno di 5 anni e ricerca di possibili cause	36
3.7	Analisi geografica sulla malnutrizione nei bambini tra i 5 ed i 19 anni e ricerca di possibili correlazioni	44
4	Power BI	48
4.1	Introduzione	48
4.1.1	ETL: Extract, Transform e Load	48
4.2	Fase di analisi	49
4.2.1	Analisi comparativa per genere	50
4.2.2	Analisi dell'impatto dell'età su sonno, stress e parametri fisiologici	51
4.2.3	Analisi dell'impatto dell'attività fisica su sonno e stress .	52
4.2.4	Analisi del benessere psicofisico in un contesto lavorativo	53

Capitolo 1

Introduzione

In questa tesina verranno presentate diverse analisi condotte su due dataset differenti: *Life Expectancy Dataset* e *Sleep Health and Lifestyle Dataset*. Il primo raccoglie dati sull'aspettativa di vita nei vari paesi, mentre il secondo si concentra sulle abitudini quotidiane e sullo stile di vita delle persone.

Le analisi sono state condotte utilizzando strumenti avanzati di data visualization e business intelligence, come **Qlik**, **Tableau** e **Power BI**,

1.1 Data Analysis Life Expectancy Dataset

Il primo dataset analizzato è Data Analysis Life Expectancy Dataset, che raccolgono informazioni sull'aspettativa di vita nei diversi stati e sui fattori che la influenzano. Il dataset è stato costruito utilizzando dati provenienti dal Global Health Observatory dell'Organizzazione Mondiale della Sanità (OMS) e dal sito delle Nazioni Unite, coprendo un periodo di 15 anni, dal 2000 al 2015, per un totale di 193 paesi.

Le variabili incluse permettono di esplorare il ruolo di diversi fattori sanitari, economici e sociali sulla longevità della popolazione. Tra questi, rientrano indicatori di mortalità, livelli di immunizzazione, condizioni economiche e variabili sociali che possono influenzare la qualità e la durata della vita. Un'attenzione particolare è stata data alla copertura vaccinale per malattie come epatite B, poliomielite e difterite, poiché il loro impatto sull'aspettativa di vita non era stato adeguatamente considerato in studi precedenti.

Il dataset consolidato contiene 22 colonne e 2938 righe, con 20 variabili predittive suddivise in quattro macro-categorie:

- **Fattori legati alla vaccinazione**
- **Fattori di mortalità**
- **Fattori economici**
- **Fattori sociali**

Di seguito, nell'elenco sono riportati i campi principali, con relative descrizioni, presenti all'interno del dataset:

1. **Country**: Nome del paese a cui si riferiscono i dati.
2. **Region**: Macro-area geografica di appartenenza del paese.
3. **Year**: Anno di riferimento della registrazione dei dati.
4. **Infant_deaths**: Numero di decessi infantili per 1.000 nati vivi.
5. **Under_five_deaths**: Numero di decessi tra i bambini sotto i cinque anni per 1.000 nati vivi.
6. **Adult_mortality**: Tasso di mortalità tra 15 e 60 anni per 1.000 abitanti.
7. **Alcohol_consumption**: Consumo medio di alcol pro capite (litri/anno).
8. **Hepatitis_B**: Percentuale della popolazione infantile vaccinata contro l'epatite B.
9. **Measles**: Numero di casi di morbillo registrati nel paese.
10. **BMI**: Indice di Massa Corporea (BMI) medio della popolazione.
11. **Polio**: Numero persone coperte dal vaccino contro il Polio.
12. **Diphtheria**: Numero persone per 1000 abitanti coperte dal vaccino contro Disfinteria.
13. **Incidents_HIV**: Numero persone per 1000 abitanti coperte dal vaccino contro HIV.
14. **GDP_per_capita**: Prodotto Interno Lordo pro capite in dollari statunitensi.
15. **Population_mln**: Popolazione totale espressa in milioni di abitanti.
16. **Thinness_ten_nineteen_years**: Numero persone per 1000 abitanti sottopeso tra i 10-19 anni.
17. **Thinness_five_nine_years**: Numero persone per 1000 abitanti sottopeso tra i 5-9 anni.
18. **Schooling**: Numero medio di anni di istruzione ricevuti.
19. **Economy_status_Developed**: 1 = Paese sviluppato, 0 = Non sviluppato.
20. **Economy_status_Developing**: 1 = Paese in via di sviluppo, 0 = Non in via di sviluppo.
21. **Life_expectancy**: Aspettativa di vita media alla nascita (anni).

22. **Country_code**: Codice a tre lettere dello stato (ISO 3166-1 alpha-3). (Aggiunta con python)
23. **Continent**: Continente di appartenenza del paese. (Aggiunta con python)
24. **Geographic_region**: Area geografica all'interno del continente. (Aggiunta con python)
25. **Thinness five 19 years**: Persone sottopeso negli anni 5-19 anni per ogni 1000 persone.

1.2 Sleep Health and Lifestyle Dataset

Il secondo dataset utilizzato per l'analisi dei dati in Power BI è Sleep Health and Lifestyle Dataset, i cui dati provengono da Google. Questo dataset, selezionato dalla piattaforma *Kaggle*, fornisce una descrizione dettagliata dello stile di vita e del benessere psico-fisico di un gruppo di persone.

Il dataset è composto da 13 colonne e circa 370 righe e ognuna di questa rappresenta una persona descritta dai seguenti attributi:

1. **Id**: identificatore univoco associato alla persona;
2. **gender**: indica il genere del soggetto (maschio o femmina);
3. **age**: rappresenta l'età dell soggetto;
4. **occupation**: indica il lavoro svolto dalla persona;
5. **sleep duration**: indica le ore di sonno giornaliere;
6. **quality of sleep**: descrive la qualità del sonno di un soggetto, che va da 1 a 10;
7. **physical activity**: rappresenta i minuti di attività fisica svolti in un giorno;
8. **stress level**: una valutazione soggettiva del livello di stress sperimentato dalla persona, che va da 1 a 10;
9. **BMI category**: indica la categoria BMI della persona (normal, overweight, obese, normal weight);
10. **blood pressure**: rappresenta la pressione del sangue data dal rapporto tra la pressione sistolica e quella diastolica;
11. **heart rate**: indica la frequenza cardiaca in battiti al minuto;
12. **sleep disorder**: indica la presenza o l'assenza di disturbi del sonno, come l'*insomnia*, ovvero quando si ha difficoltà ad addormentarsi e la *sleep apnea*, che prevede interruzioni del sonno dovute a difficoltà respiratorie;
13. **daily steps**: rappresenta il numero di passi giornalieri compiuti.

Capitolo 2

Qlik

2.1 Introduzione

Qlik è una società di software specializzata in soluzioni di Business Intelligence (BI) e Data Analytics. Fondata in Svezia nel 1993 e oggi con sede negli Stati Uniti, Qlik è nota per il suo approccio innovativo all'analisi dei dati, basato su un motore associativo che permette agli utenti di esplorare le informazioni in modo intuitivo e interattivo.

- **QlikView:** una piattaforma di Guided Analytics, pensata per fornire analisi guidate attraverso dashboard e report preconfigurati. Le applicazioni vengono sviluppate da analisti o sviluppatori e permettono agli utenti di interagire con i dati, ma senza la possibilità di creare o modificare visualizzazioni in autonomia.
- **Qlik sense:** una piattaforma di Self-Service Analytics, progettata per offrire agli utenti la libertà di esplorare i dati in modo autonomo e creare dashboard interattive personalizzate, senza necessità di competenze di programmazione.

A differenza di molte altre soluzioni BI basate su database relazionali e query SQL, Qlik utilizza un motore associativo, che elimina le limitazioni imposte dalle strutture rigide delle query predefinite. Questo consente agli utenti di navigare liberamente tra i dati, scoprendo connessioni e insight senza dover ricorrere a complesse riscritture di query.

2.2 Caricamento dei dati

Il processo di caricamento dei dati in Qlik è un passaggio essenziale che permette di importare informazioni da diverse fonti, tra cui file CSV. Prima dell'importazione, il file deve essere preparato con un'intestazione chiara e dati coerenti. Una volta aperto l'Editor di Script in Qlik, si seleziona il file .CSV da caricare

e si configurano le impostazioni necessarie, come il delimitatore e il tipo di carattere. Dopo l'esecuzione dello script, Qlik importa i dati e genera un modello associativo, che facilita l'analisi interattiva.

Nel caso specifico di questa tesina, il file .CSV è stato caricato direttamente, senza utilizzare la funzione "Gestione Dati" per combinare più file, poiché tutte le informazioni necessarie erano già contenute in un unico dataset. Questo approccio ha reso il processo più semplice e rapido. Inoltre, nella fase di Gestione Dati, le celle vuote sono state sostituite con il valore NULL, garantendo una gestione più accurata dei dati mancanti

2.3 Data Analysis: Life Expectancy (WHO) Fixed

Dopo il caricamento dei dati, l'analisi del dataset è stata condotta attraverso l'uso delle dashboard, che offrono strumenti avanzati per la creazione di grafici, l'applicazione di filtri e l'esplorazione approfondita delle informazioni. Queste dashboard possono essere personalizzate con nomi distintivi e organizzate in base agli argomenti di analisi, rendendo la navigazione tra i diversi set di dati più intuitiva ed efficiente.

Nelle sezioni seguenti verranno analizzate le dashboard sviluppate a partire dal dataset "Life Expectancy (WHO) Fixed", con una panoramica sulle funzionalità, i grafici e le analisi specifiche presenti in ciascuna di esse.

- **Informazioni generali.**
- **Analisi GDP e Vita**
- **Analisi popolazione e morte infatile e adulta**
- **Analisi tra BMI e morte infantile**
- **Analisi tra BMI e morte adulta**
- **Relazione tra Schooling e Adults Deaths**
- **Relazione tra Schooling e Infants Deaths**

Dashboard 1: Informazioni generali.

La dashboard mostrata in **Figura 2.1** offre una panoramica completa delle informazioni presenti nel dataset. Essa evidenzia il numero di paesi inclusi, i continenti coinvolti, la media della popolazione per paese nel corso degli anni, il PIL pro capite medio, la media degli anni scolastici e il BMI medio. Sono inoltre presenti i principali filtri, che possono essere applicati anche nelle dashboard successive (Anno, Paese, Continente e Regioni geografiche). Questi dati saranno analizzati più nel dettaglio nelle dashboard successive. Al variare dei filtri, i grafici si aggiorneranno automaticamente per riflettere le nuove variabili selezionate.

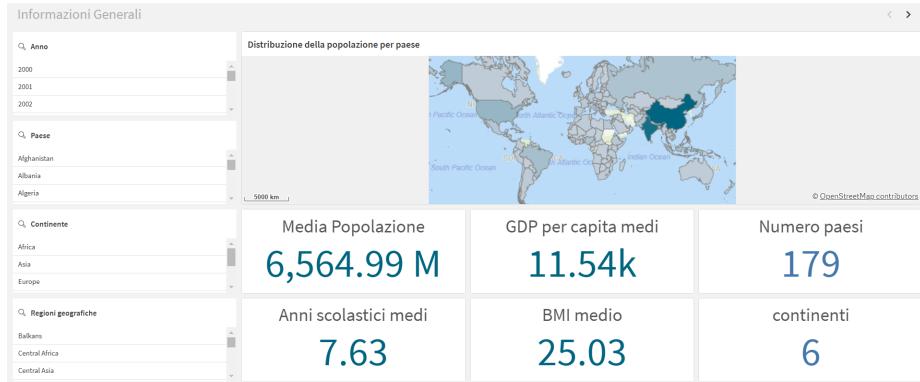


Figura 2.1

Filtri

I filtri rappresentano uno strumento fondamentale per affinare e personalizzare l'analisi dei dati. Nelle seguenti sezioni vengono illustrati due principali tipi di filtri, che possono essere applicati per focalizzarsi su vari aspetti del dataset:

- **Arearie Geografiche:** questo filtro consente di selezionare i dati in base a vari livelli geografici, come il paese, il continente o una regione specifica. Un esempio di applicazione di questo filtro è visibile nella **Figura 2.2**, dove è stato selezionato il filtro per il continente, con un focus particolare sull'Europa; la dashboard mostra la media nel lasso di anni considerato.

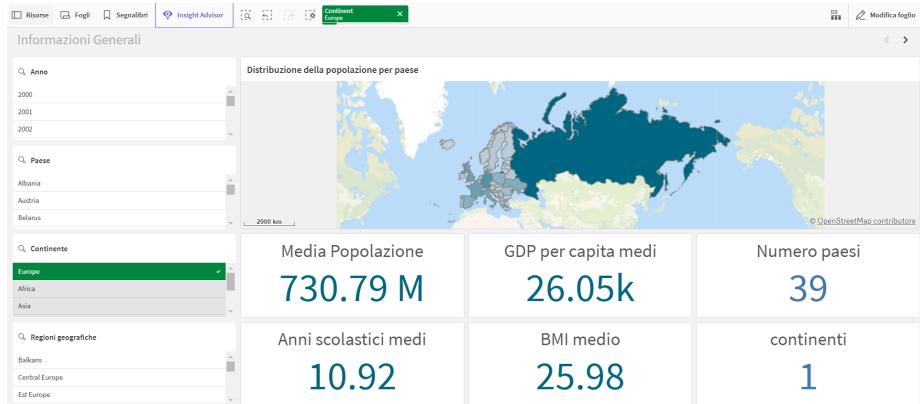


Figura 2.2

- **Anno:** questo filtro permette di osservare l'evoluzione dei dati nel corso degli anni. Nella **Figura 2.3**, viene mostrato come i dati specifici si presentano per l'anno 2010, fornendo una visione dettagliata delle tendenze in quel periodo.

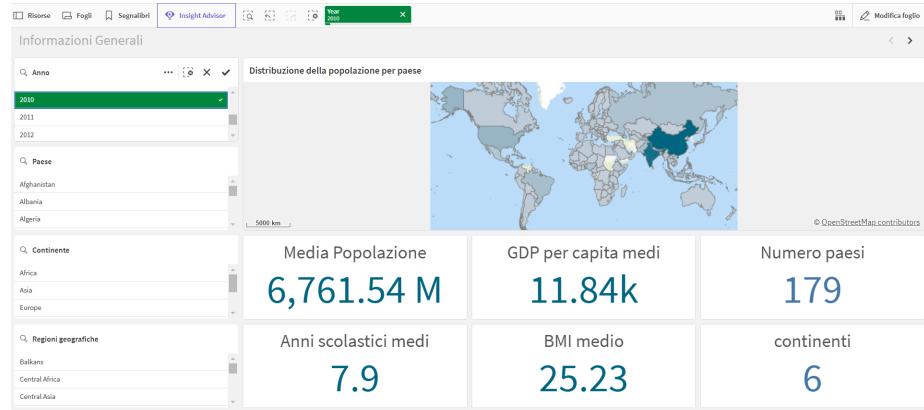


Figura 2.3

Oltre all’uso singolo dei filtri, è anche possibile combinarli per ottenere analisi più mirate e approfondite. Ad esempio, applicando contemporaneamente il filtro per continente e per anno, si può ottenere una visione più specifica e rilevante di come i dati si distribuiscono in determinate aree geografiche e in periodi temporali specifici.

I filtri non solo rendono la navigazione tra i dati più fluida, ma offrono anche la possibilità di identificare pattern e tendenze che potrebbero non emergere da un’analisi superficiale. Di seguito, presentiamo un esempio che illustra l’evoluzione del continente africano, a partire dal 2000 **Figura 2.4**, primo anno considerato, fino al 2015, ultimo anno per cui sono disponibili i dati **Figura 2.5**.

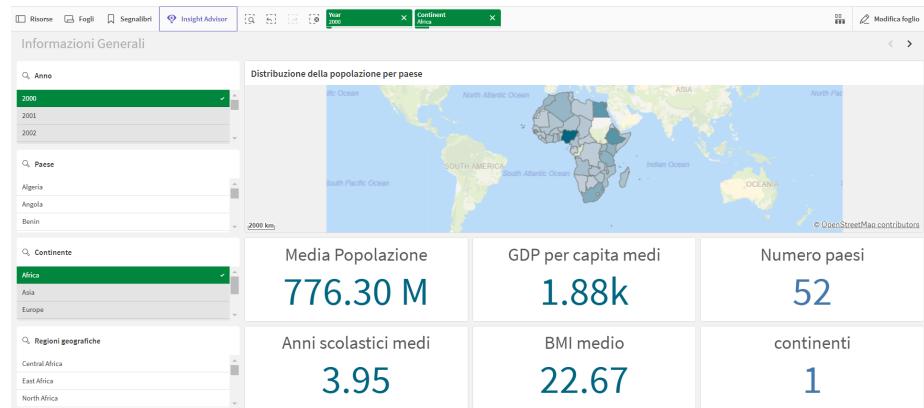


Figura 2.4

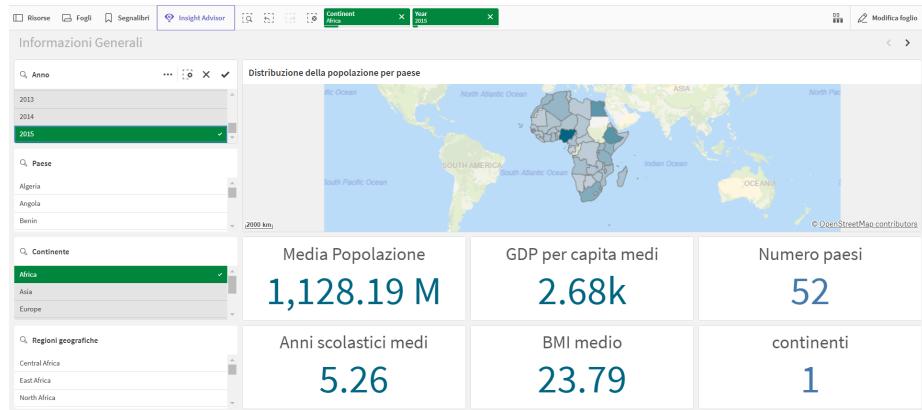


Figura 2.5

Come si può osservare per il continente africano, considerando questi due anni, si nota un aumento di tutti i fattori analizzati, tra cui i più significativi sono la popolazione, gli anni scolastici, il PIL pro capite e il BMI. Una tendenza simile è visibile anche negli altri continenti, sebbene con aumenti meno rilevanti. Questi aspetti saranno successivamente analizzati nel dettaglio per comprendere il loro impatto su parametri come l'aspettativa di vita, la mortalità infantile e la mortalità adulta.

Dashboard 2: Analisi GDP e Vita.

La **Dashboard 2** ha l'obiettivo di analizzare la relazione tra il GDP pro capite medio della popolazione e tre parametri chiave: aspettativa di vita, mortalità infantile e mortalità adulta.

Sono considerati tutti i paesi, con una distinzione cromatica che rappresenta i diversi continenti, come illustrato nella **Figura 2.6**. La dashboard è suddivisa in cinque grafici principali, ognuno dei quali utilizza una specifica tipologia di visualizzazione:

- Relazione tra GDP pro capite e aspettativa di vita – grafico a dispersione;
- Relazione tra GDP pro capite e mortalità infantile nei vari stati – grafico a dispersione;
- Relazione tra GDP pro capite e mortalità adulta – grafico a dispersione;
- Andamento del GDP nei diversi continenti per anno – grafico a linee;
- Evoluzione dell'aspettativa di vita nei vari continenti nel corso degli anni – grafico a linee.

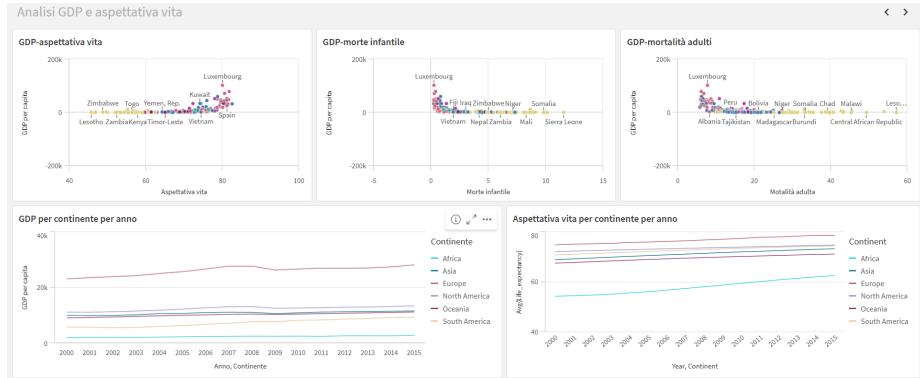


Figura 2.6

Dall'analisi di questi grafici emerge chiaramente una relazione tra il GDP pro capite e gli altri aspetti cruciali della vita, come l'aspettativa di vita, la mortalità infantile e la mortalità adulta. In particolare, osserviamo che i paesi con un GDP pro capite inferiore si collocano generalmente nella parte meno favorevole per tutti e tre i parametri analizzati. In altre parole, i paesi con minori risorse economiche tendono ad avere aspettative di vita più basse e livelli più alti di mortalità infantile e adulta. Al contrario, i paesi con un GDP pro capite maggiore si trovano nella parte superiore dei grafici, con valori migliori per l'aspettativa di vita e tassi di mortalità infantile e adulta più bassi.

Questa relazione è ancora più evidente se consideriamo i paesi per continente. Nei grafici di dispersione, i paesi con un GDP pro capite più basso, come quelli africani (individuati dal colore giallo), tendono a presentare i valori peggiori per aspettativa di vita e mortalità. Al contrario, i paesi europei, rappresentati dal colore rosso, mostrano parametri migliori, evidenziando la differenza tra continenti sviluppati e in via di sviluppo, come mostrato in **Figura 2.7** in cui si evidenziano il continente europeo e quello africano.

I due grafici a linee mostrano l'andamento del GDP pro capite e dell'aspettativa di vita per ogni continente nel periodo dal 2000 al 2015, con i dati distinti per continente attraverso una codifica cromatica. Il primo grafico evidenzia l'aumento del GDP, mentre il secondo illustra l'evoluzione dell'aspettativa di vita nel tempo.

Si osserva una chiara relazione tra l'incremento del GDP e il miglioramento dell'aspettativa di vita in tutti i continenti. Tuttavia, in Europa, dove il GDP pro capite è già elevato, l'incremento dell'aspettativa di vita risulta relativamente contenuto. Al contrario, in Africa, nonostante l'incremento modesto del GDP, si registra un miglioramento più significativo nell'aspettativa di vita. Questo suggerisce che, a livelli di GDP più bassi, anche un aumento moderato del reddito pro capite può avere un impatto più significativo sull'aspettativa di vita rispetto ai paesi con un GDP già elevato.

Inoltre, si può osservare che, nonostante la crisi economica del 2008, l'aspettativa di vita nei vari continenti ha continuato a crescere in modo costante.

Questo potrebbe sembrare contortivo, ma suggerisce che, sebbene un miglioramento economico possa influire positivamente sull'aspettativa di vita, una recessione momentanea non ha un impatto immediato su di essa. In altre parole, mentre i cambiamenti economici a lungo termine possono condizionare la qualità della vita, un rallentamento temporaneo non sembra alterare significativamente l'andamento generale dell'aspettativa di vita, come mostrato in **Figura 3.5**.

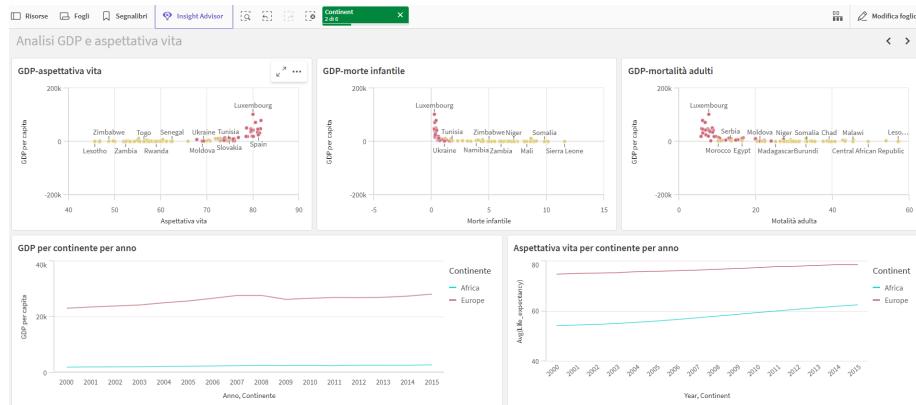


Figura 2.7

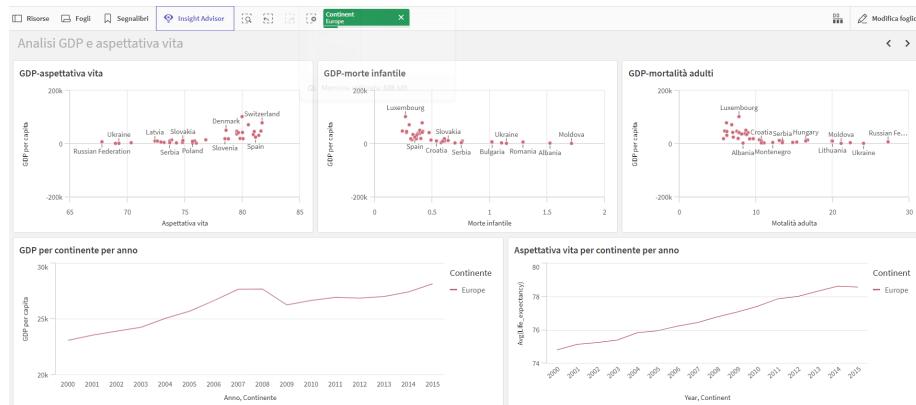


Figura 2.8

Inoltre selezionando un unico paese possiamo vedere i dati relativi ad esso e come tale siano sviluppati nel corso degli anni

Dashboard 3: Analisi popolazione e morte infattile e adulta.

Nella **Dashboard 3** viene esplorata la possibile relazione tra la popolazione di un paese e le relative percentuali di mortalità infantile e adulta. Questa

dashboard è composta da sei grafici **Figura 2.9**, descritti di seguito:

- Un istogramma della mortalità infantile, che illustra la distribuzione percentuale dei decessi neonatali nei vari paesi, con una distinzione cromatica per continente.
- Un istogramma della mortalità adulta, che mette in evidenza il tasso di decessi tra la popolazione adulta nei diversi paesi, anch'esso suddiviso per continente con codifica cromatica.
- Un diagramma a torta della popolazione mondiale, che visualizza la ripartizione della popolazione tra i vari stati.
- Un grafico a linee che mostra l'andamento della crescita della popolazione nel tempo, suddiviso per continente.
- Un istogramma che rappresenta il numero assoluto di decessi per mortalità adulta, evidenziando la quantità di decessi tra la popolazione adulta nei vari paesi.
- Un istogramma che visualizza il numero assoluto di decessi per mortalità infantile, mostrando la distribuzione numerica dei decessi neonatali nei vari paesi.

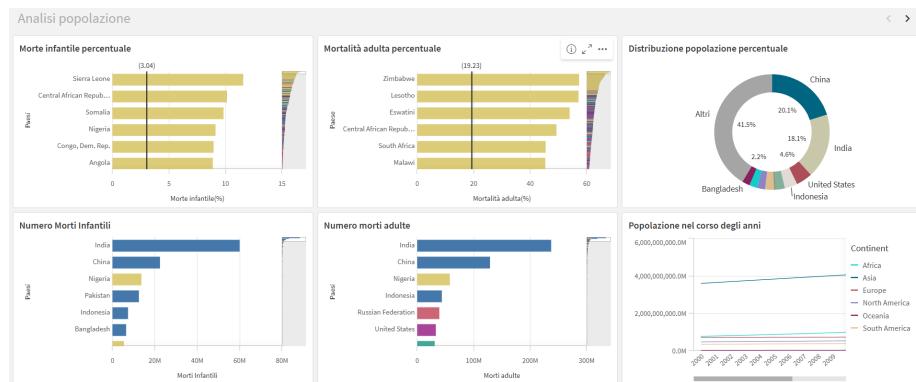


Figura 2.9

Dall'analisi dei dati emerge che il numero di abitanti di un paese non ha un impatto diretto sulle percentuali di mortalità, risultando un fattore relativamente irrilevante in relazione ai tassi di mortalità infantile e adulta. Nei grafici che mostrano la percentuale di mortalità infantile e adulta in relazione alla popolazione, i paesi africani sono più rappresentati, il che suggerisce una concentrazione più alta di mortalità in questi paesi, ma non implica necessariamente una correlazione diretta con la dimensione della popolazione. Tuttavia,

quando si considera il numero assoluto di decessi, si evidenzia un notevole distacco dei due paesi asiatici, India e Cina, a causa della loro enorme popolazione. Infatti, il numero totale di decessi in questi due paesi supera quello dell'intero continente africano, il che dimostra come la densità demografica possa influire notevolmente sui numeri assoluti, ma non necessariamente sui tassi percentuali di mortalità.

Un aspetto rilevante che emerge dall'analisi dei dati è che i paesi con i tassi di mortalità infantile e adulta più elevati si concentrano principalmente nel continente africano. In molti di questi paesi, i valori dei tassi di mortalità superano la media globale, come si osserva chiaramente nei grafici, dove le barre gialle eccedono la linea che rappresenta la media globale. Questo evidenzia le difficoltà persistenti in Africa nel migliorare la salute pubblica e l'accesso ai servizi sanitari, nonostante gli sforzi in atto. Tuttavia, questa tendenza non è altrettanto evidente nei grafici che mostrano il numero assoluto di decessi infantili e adulti, dove la distribuzione appare più articolata e complessa, con una maggiore variabilità tra i vari paesi e continenti.

Un confronto particolarmente interessante emerge tra due continenti che si pongono agli estremi opposti in termini di indicatori: l'Africa e l'Europa. Nei grafici relativi ai tassi di mortalità, i paesi africani occupano prevalentemente le posizioni più alte, mentre, quando si analizzano i numeri effettivi dei decessi, il continente africano emerge come predominante, nonostante la sua popolazione sia inferiore a quella di molti stati europei. Questo fenomeno suggerisce che, sebbene i paesi europei abbiano una popolazione numericamente maggiore rispetto a molte nazioni africane, i tassi di mortalità in Africa continuano a essere drammaticamente più elevati. Questo divario si manifesta chiaramente nei grafici e rappresenta una sfida significativa per le politiche sanitarie africane, come ben illustrato nella **Figura 2.10**.

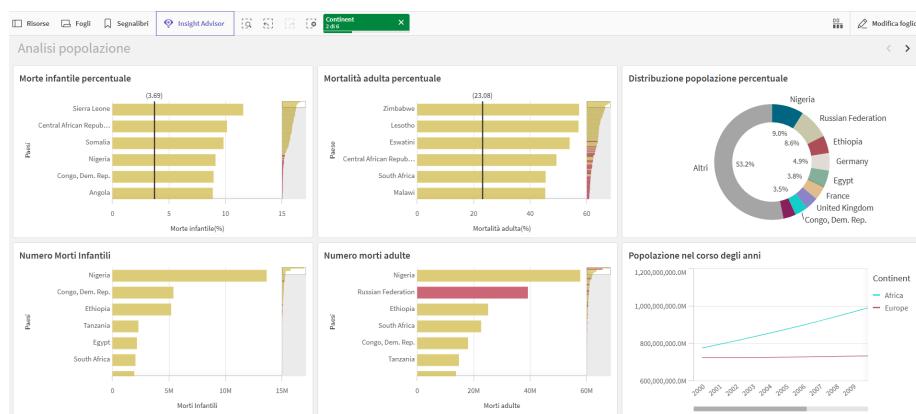


Figura 2.10

Successivamente, presentiamo un'analisi focalizzata sul continente africano, in quanto rappresenta una delle aree geografiche che subisce maggiormente gli

effetti dei parametri in esame. Questa analisi si concentra sui dati relativi agli anni 2005, **Figura 2.11**, e 2015, **Figura 2.12**, in modo da confrontare l'evoluzione delle principali tendenze in un arco temporale di dieci anni. Analizzando i risultati, si osserva che la percentuale di mortalità infantile ha subito una riduzione significativa, passando dal 7,5% nel 2005 al 4,1% nel 2015. Questo dato suggerisce un miglioramento delle condizioni sanitarie e dei sistemi di assistenza per i neonati in molte nazioni africane. Allo stesso modo, la percentuale di mortalità adulta è diminuita dal 36,95% al 26,22% nello stesso periodo, indicando un miglioramento delle condizioni di vita e dell'accesso ai servizi sanitari per la popolazione adulta.

Tuttavia, è importante sottolineare che, nonostante queste riduzioni percentuali, il numero assoluto di decessi, sia infantili che adulti, rimane sostanzialmente invariato. Questo fenomeno può essere spiegato dal forte aumento della popolazione del continente, che è passata da circa 800 milioni a oltre 1,1 miliardi di abitanti nel periodo considerato. Il drastico incremento demografico ha infatti controbilanciato i miglioramenti nelle percentuali di mortalità, mantenendo invariato il numero complessivo di decessi.

Nonostante queste riduzioni percentuali nei tassi di mortalità, il continente africano continua a presentare i tassi di mortalità più elevati sia per i neonati che per gli adulti. Questo evidenzia come, nonostante i progressi compiuti, le sfide sanitarie e sociali del continente restino significative. La crescita della popolazione, purtroppo, continua a essere un fattore che amplifica il numero totale di decessi, facendo sì che l'Africa rimanga una delle regioni con i parametri più critici a livello globale.

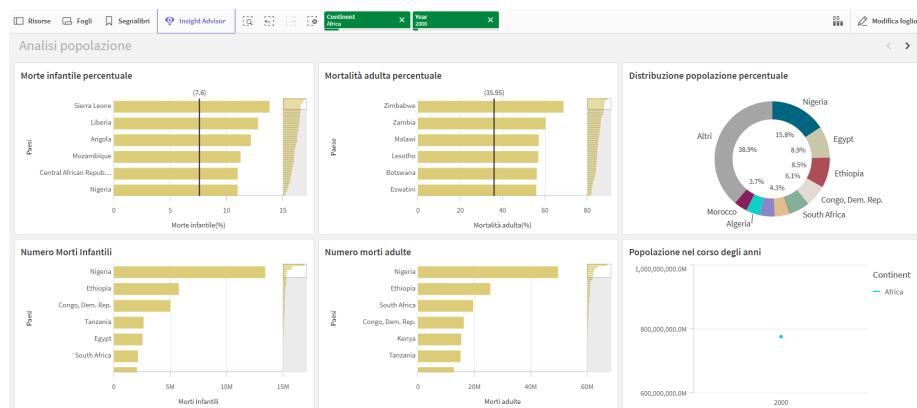


Figura 2.11

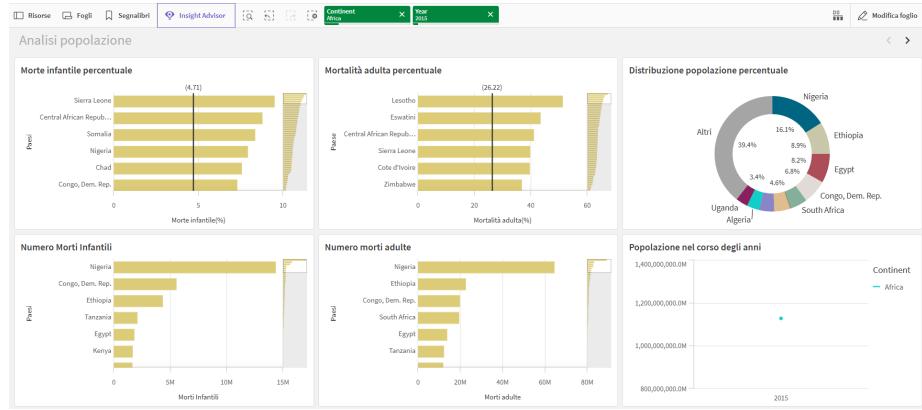


Figura 2.12

In sintesi, l'analisi dei dati evidenzia come la mortalità infantile e adulta sia un fenomeno complesso che non può essere compreso esclusivamente in relazione alla popolazione di un paese. Sebbene l'Africa presenti i tassi di mortalità più elevati, il numero assoluto di decessi è significativamente influenzato dalla densità demografica, con paesi come India e Cina che, pur avendo tassi inferiori, mostrano numeri assoluti di decessi che superano quelli dell'intero continente africano. La disparità tra i continenti emerge anche nei grafici relativi ai tassi di mortalità, dove l'Africa domina le posizioni più alte. Questi dati suggeriscono la necessità di interventi mirati e differenziati, specialmente in Africa, dove l'accesso ai servizi sanitari e la qualità delle politiche sanitarie rimangono sfide urgenti nonostante i miglioramenti ottenuti negli ultimi anni.

Dashboard 4: Analisi tra BMI e mortalità infantile.

L'Indice di Massa Corporea (IMC) nei bambini è un importante indicatore dello stato nutrizionale e si correla strettamente al tasso di mortalità infantile. Come si osserva nei grafici, in base alle regioni, esiste una chiara relazione tra IMC e mortalità infantile. Esaminiamo i principali esempi:

- **Asia:** Nel continente asiatico si osserva un IMC relativamente basso, ma una tendenza crescente nella mortalità infantile. Questa situazione è particolarmente evidente nei paesi in via di sviluppo del sud-est asiatico, come Laos, Myanmar, Cambogia e Vietnam, dove si verificano frequentemente problemi climatici e una dieta non equilibrata.
- **Africa:** L'Africa, in particolare nella regione centrale e occidentale, presenta un IMC molto basso. Questa condizione è legata a fattori come guerre, difficoltà climatiche e malattie. In particolare, la regione centrale del continente, che comprende la giungla del Congo, è caratterizzata da una fauna e flora ostile che rende difficile la creazione di paesi stabili. Nel-

la regione occidentale, il Sahara influisce negativamente sull'agricoltura, rendendo difficile garantire una nutrizione adeguata.

- **Europa, America del Sud e Oceania:** Queste tre regioni mostrano dati simili. Quando l'IMC si mantiene nei valori raccomandabili, il tasso di mortalità infantile tende a essere basso.
- **America del Nord:** Qui, purtroppo, si osserva una tendenza diversa rispetto al resto del mondo. Sebbene l'IMC si trovi nei valori raccomandabili, il tasso di mortalità infantile non rispecchia tale situazione. Le cause principali includono una scarsa qualità del cibo, la mancanza di consapevolezza sulla salute e l'influenza del sistema sanitario privato.

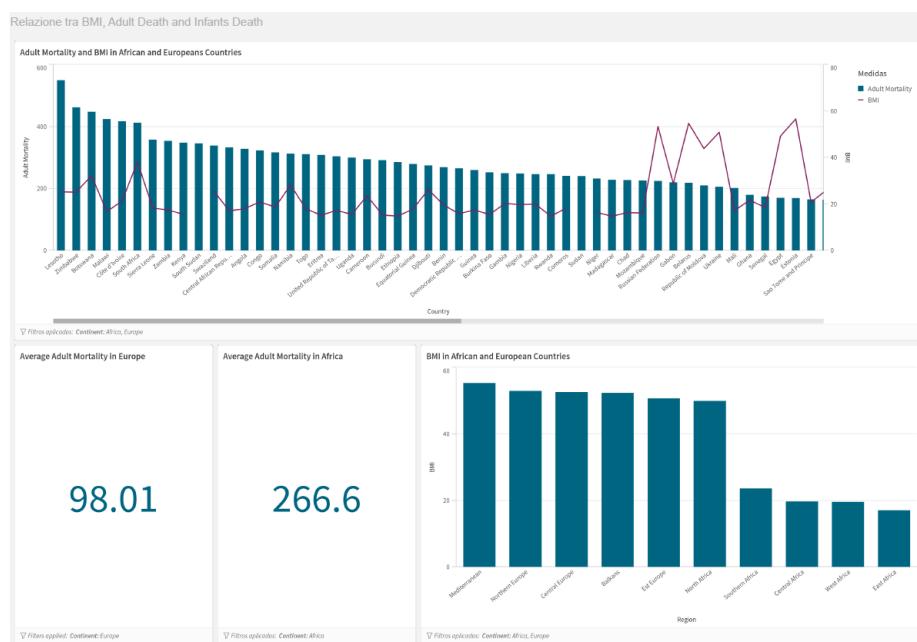


Figura 2.13

Dashboard 5: Analisi tra BMI e mortalità adulta.

- **America:** Negli Stati Uniti e in alcune regioni dell'America Latina, l'obesità è diventata una problematica di salute pubblica. Un IMC elevato è correlato a una maggiore mortalità per malattie cardiovascolari, diabete di tipo 2 e cancro. Tuttavia, in alcune aree rurali dell'America Latina, il sottopeso rimane un fattore di rischio a causa della malnutrizione.
- **Europa:** In Europa, la situazione varia notevolmente da una regione all'altra. Nei paesi dell'Europa occidentale, l'aumento del sovrappeso e del-

l'obesità porta a un incremento delle malattie metaboliche e cardiovascolari. Nell'Europa dell'Est, le popolazioni vulnerabili, esposte al sottopeso e alla malnutrizione, hanno una speranza di vita più bassa.

- **Africa:** In Africa, esiste una doppia sfida: da un lato, la malnutrizione e il sottopeso sono fattori significativi di mortalità; dall'altro, nelle aree urbane, l'obesità sta diventando un problema crescente, aumentando il rischio di malattie croniche. In generale, la malnutrizione, sia per eccesso che per difetto, rimane una grave questione di salute pubblica.
- **Asia:** In Asia, le differenze regionali sono notevoli. Nel sud e nel sud-est asiatico, il sottopeso è ancora un fattore di rischio, soprattutto per donne e bambini. In Asia orientale, paesi come Cina e Giappone stanno affrontando un aumento dell'obesità, con un conseguente incremento delle malattie cardiovascolari e metaboliche.
- **Oceania:** In Oceania, paesi come Australia e Nuova Zelanda presentano tassi elevati di obesità, con un impatto negativo sulla mortalità legata a malattie croniche. Tuttavia, in alcune isole del Pacifico, la malnutrizione rimane un problema prevalente in alcune popolazioni.

In conclusione, la relazione tra IMC e mortalità varia a seconda del continente e del contesto socioeconomico. Tuttavia, sia il sottopeso che l'obesità aumentano il rischio di malattie e riducono l'aspettativa di vita, rendendo essenziale un bilanciamento nutrizionale adeguato.

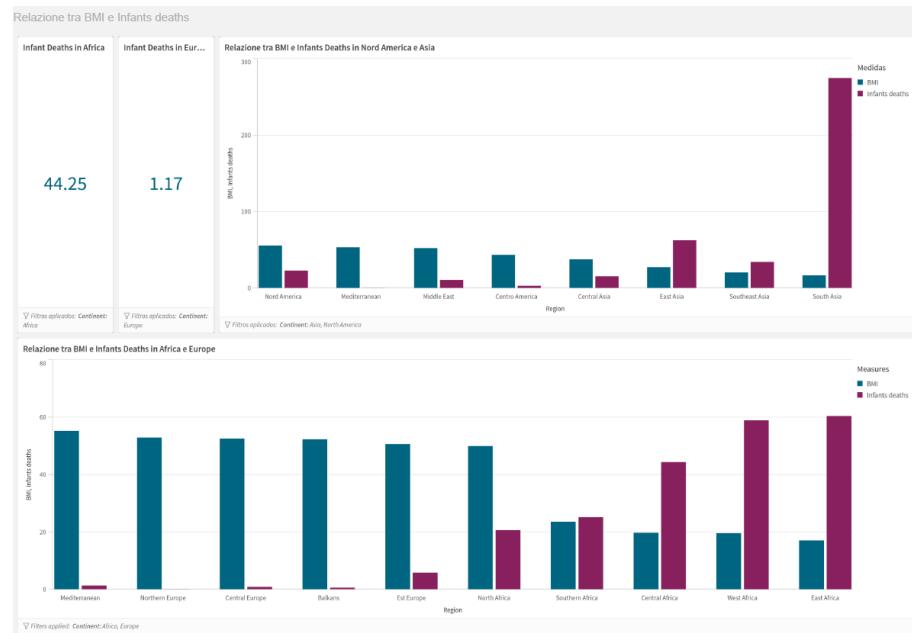


Figura 2.14

Dashboard 6: Relazione tra Schooling e Adults Deaths

Relazione tra Schooling e Adults Deaths La mortalità adulta e il tasso di scolarizzazione sono correlati, anche se il loro impatto varia significativamente tra le diverse regioni del mondo. In Africa, dove i livelli di scolarizzazione sono più bassi rispetto all'Europa, i tassi di mortalità adulta rimangono elevati a causa di fattori come l'accesso limitato alla sanità, la povertà e la diffusione di malattie. Al contrario, in Europa, dove l'istruzione è accessibile e obbligatoria, l'aspettativa di vita è più alta e la mortalità adulta è diminuita considerevolmente. Ma non si vede una tendenza così forte come rispetto alla mortalità infantile che si può vedere nel grafico.

In Africa, la mancanza di istruzione influisce direttamente sull'elevata mortalità adulta. Molte persone non hanno accesso a informazioni sulla prevenzione delle malattie, rendendole più vulnerabili a infezioni come malaria e tubercolosi. Inoltre, il basso livello di scolarizzazione limita le opportunità lavorative, perpetuando la povertà e rendendo difficile l'accesso a un'alimentazione adeguata e a servizi sanitari di qualità. L'instabilità politica e i conflitti armati in alcune regioni aggravano ulteriormente la situazione, incidendo sia sull'istruzione che sui sistemi sanitari.

In Europa, l'istruzione ha svolto un ruolo fondamentale nella riduzione della mortalità adulta. Un livello di scolarizzazione più alto è associato a migliori abitudini di salute, a un minor consumo di tabacco e alcol e a un più ampio accesso ai servizi medici. Inoltre, le economie sviluppate e i sistemi di welfare europei garantiscono che le persone istruite abbiano accesso a lavori meglio retribuiti, consentendo loro di mantenere una qualità della vita più elevata e di ricevere cure mediche adeguate.

Le differenze tra Africa ed Europa dimostrano come l'istruzione abbia un impatto diretto sulla salute e sulla longevità della popolazione. Mentre in Europa la scolarizzazione ha contribuito a migliorare la salute pubblica e a ridurre le malattie prevenibili, in Africa garantire l'accesso all'istruzione rimane una sfida, soprattutto nelle aree rurali ed emarginate. Per ridurre la mortalità adulta in Africa, è essenziale investire nell'istruzione, rafforzare i sistemi sanitari e garantire l'accesso alle informazioni sulla prevenzione delle malattie e sulla promozione della salute.

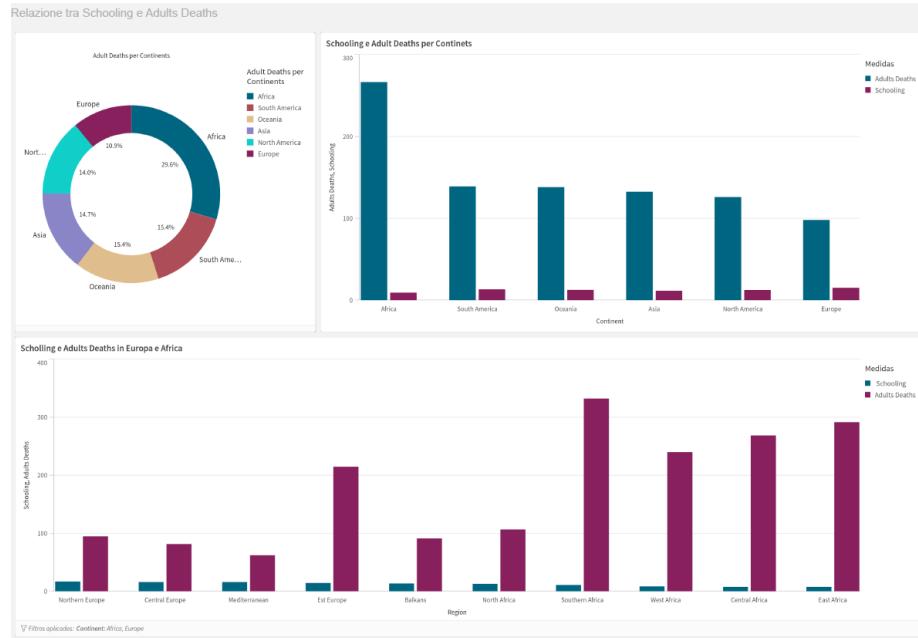


Figura 2.15

Dashboard 7: Relazione tra Schooling e Infants Deaths

In questi grafici si vede come esiste una chiara relazione tra la tassa di scolarizzazione e i morti infantile. La mortalità infantile e il tasso di scolarizzazione sono strettamente interconnessi, anche se il loro impatto varia tra le diverse regioni del mondo. In Africa, dove i tassi di scolarizzazione sono più bassi rispetto all'Europa, la mortalità infantile rappresenta ancora una sfida importante. Al contrario, in Europa, dove l'istruzione è più accessibile e obbligatoria, i tassi di mortalità infantile sono diminuiti significativamente negli ultimi decenni.

In Africa, la mancanza di accesso all'istruzione, soprattutto per le ragazze, contribuisce all'elevata mortalità infantile. Molte donne non ricevono informazioni adeguate sulla salute materno-infantile, aumentando così il rischio di complicazioni durante la gravidanza e i primi anni di vita del bambino. Inoltre, la povertà, l'instabilità politica e la carenza di infrastrutture sanitarie aggravano questa situazione. Secondo i dati degli organismi internazionali, i paesi africani con i tassi di scolarizzazione più bassi registrano spesso i livelli più alti di mortalità infantile.

D'altra parte, in Europa, l'istruzione universale e l'accesso a servizi sanitari di qualità sono stati fondamentali per ridurre la mortalità infantile a livelli molto bassi. L'istruzione obbligatoria, insieme a campagne di sensibilizzazione sulla salute materno-infantile, ha permesso alle donne europee di acquisire maggiore conoscenza su pratiche di igiene, nutrizione e prevenzione delle malattie. Inoltre,

i sistemi sanitari europei garantiscono cure mediche durante la gravidanza e l'infanzia, contribuendo significativamente alla sopravvivenza dei bambini.

Le differenze tra Africa ed Europa evidenziano come l'istruzione sia un fattore determinante nella riduzione della mortalità infantile. Mentre in Europa la scolarizzazione ha aiutato a eliminare le barriere di accesso alla salute, in Africa rimane una sfida, soprattutto nelle aree rurali e più emarginate. Per ridurre la mortalità infantile in Africa, è fondamentale investire nell'istruzione, garantire un accesso equo alla scuola per bambine e bambini e rafforzare i sistemi sanitari pubblici.

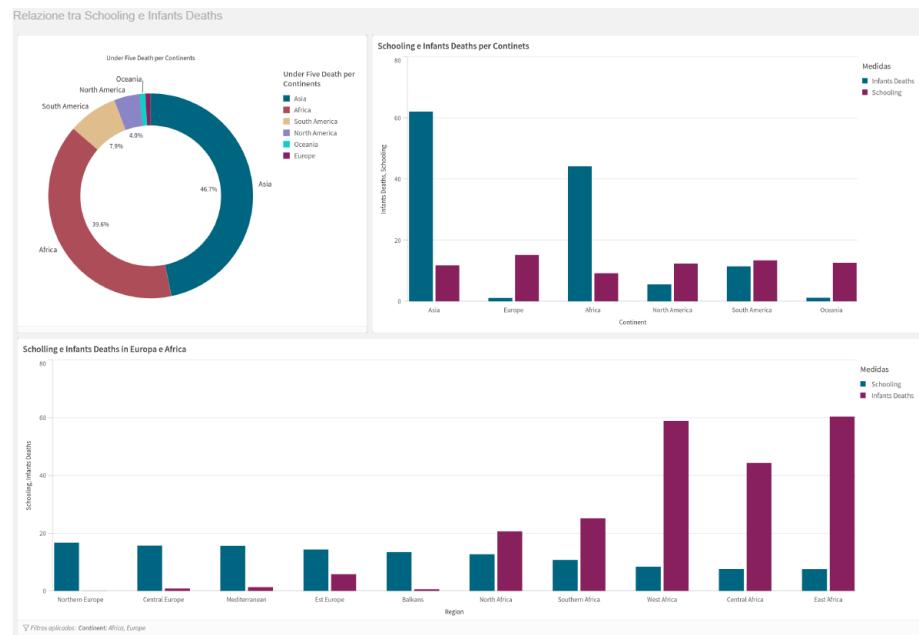


Figura 2.16

Capitolo 3

Tableau

3.1 Introduzione

Tableau è una potente piattaforma di visualizzazione e analisi dei dati, progettata per rendere l'esplorazione delle informazioni intuitiva e accessibile anche a chi non ha competenze avanzate di programmazione. Nato nel 2003 da un team di ricercatori della Stanford University, tra cui Chris Stolte, Christian Chabot e Pat Hanrahan, il software è stato sviluppato con l'obiettivo di semplificare il processo di trasformazione dei dati grezzi in grafici e dashboard interattivi.

Alla base di Tableau c'è l'innovativo linguaggio VizQL (Visual Query Language), che permette di convertire automaticamente complesse interrogazioni in rappresentazioni visive chiare e immediate. Grazie a questa tecnologia, gli utenti possono analizzare grandi quantità di dati provenienti da diverse fonti, applicando filtri, evidenziazioni, proiezioni e tendenze per estrarre insight significativi.

La sua interfaccia intuitiva e le avanzate capacità di analisi visiva hanno contribuito alla rapida diffusione di Tableau in numerosi settori, trasformandolo in uno strumento indispensabile per aziende, analisti e manager. Con Tableau, la business intelligence diventa più accessibile, consentendo di prendere decisioni strategiche basate su dati concreti, senza necessità di conoscenze tecniche approfondite.

3.2 Gestione dei dati

In Tableau, l'analisi dei dati si concentra principalmente sulla capacità di connettersi a diverse fonti di dati e di manipolare le informazioni in modo semplice grazie alle funzionalità di drag-and-drop. Questo approccio consente di trasformare i dati grezzi in visualizzazioni grafiche personalizzate, facilmente interpretabili.

Il processo inizia con il collegamento a una o più fonti di dati, che possono includere database, file Excel, fonti web o file CSV (come i dataset utilizzati in

questa tesina). Una volta stabilita la connessione, Tableau permette di esplorare, pulire e preparare i dati direttamente all'interno dell'interfaccia. Ad esempio, è possibile rimuovere duplicati, gestire valori nulli, creare campi calcolati e unire set di dati differenti, rendendo l'intero processo di analisi molto più dinamico e accessibile.

3.3 Rappresentazioni dei dati

Una delle caratteristiche che ha contribuito al successo di Tableau è la sua capacità di visualizzare i dati attraverso una vasta gamma di opzioni grafiche. Dalle semplici curve di tendenze e grafici a barre fino alle mappe geografiche interattive, il software offre molteplici modalità per rappresentare visivamente le informazioni, sebbene la scelta della visualizzazione dipenda dalla tipologia dei dati.

Un altro aspetto distintivo di Tableau è la sua interattività. Ogni grafico generato può essere esplorato in modo dinamico, grazie alle Dashboard, che sono pagine composte da più visualizzazioni unite in un'unica vista. Le Dashboard consentono agli utenti di filtrare e manipolare i dati in tempo reale per ottenere informazioni più approfondite.

Inoltre, Tableau consente la creazione di storytelling attraverso le story, ovvero sequenze di grafici e dashboard strutturate per raccontare una narrazione visiva basata sui dati. Questa funzionalità aiuta a guidare il pubblico nel processo di esplorazione, enfatizzando le tendenze e le scoperte principali.

Infine, come accennato nella sezione precedente, uno degli elementi che ha reso Tableau un software di successo è l'integrazione di algoritmi statistici per la generazione di previsioni. Funzioni come la regressione lineare permettono di proiettare tendenze future sulla base dei dati storici, aiutando gli utenti a interpretare come evolveranno determinate variabili nel tempo.

3.4 Data Analysis Life Expectancy Dataset

Dopo aver completato il caricamento dei dati, si è iniziata la fase di analisi del dataset. Per dare una visione globale e di chiara rappresentazione si è deciso di dividere le analisi in vari fenomeni. Per ogni fenomeno analizzato sono stati realizzati dei **Fogli di lavoro** o delle **Dashboard**. I fenomeni analizzati sono e seguenti:

- **Analisi sulla copertura del vaccino contro il morbillo e correlazioni con la mortalità nei bambini**
 - I. **Dashboard 1:** Analisi dell'andamento della copertura vaccinale contro il morbillo nei continenti
 - II. **Dashboard 2:** Analisi dell'andamento della copertura del vaccino contro il morbillo nelle regioni

- III. **Dashboard 3:** Analisi dell'andamento della copertura del vaccino contro il morbillo nelle regioni
- IV. **Dashboard 4:** Analisi della relazione tra copertura del vaccino contro il morbillo e morti nei bambini
- **Analisi sul tasso di mortalità nei bambini con meno di 5 anni e ricerca di possibili cause**
 - I. **Dashboard 5:** Analisi della distribuzione geografica delle morti nei bambini con meno di 5 anni
 - II. **Dashboard 6:** Analisi della correlazione tra morti nei bambini con meno di 5 anni e GDP per capita
 - III. **Dashboard 7:** Analisi della correlazione tra le coperture vaccinali e le morti nei bambini con meno di 5 anni
- **Analisi geografica sulla malnutrizione nei bambini tra i 5 ed i 19 anni e ricerca di possibili correlazioni**
 - I. **Dashboard 8:** Analisi geografica della malnutrizione nei continenti, nelle regioni e negli stati

3.5 Analisi sulla copertura del vaccino contro il morbillo e correlazioni con la mortalità nei bambini

L'analisi dei dati prende avvio esaminando la copertura vaccinale contro il morbillo a livello globale, considerando il ruolo cruciale del vaccino nella prevenzione di questa malattia altamente contagiosa. Ogni anno, il morbillo colpisce oltre 9 milioni di persone, causando più di 100.000 decessi. L'Organizzazione Mondiale della Sanità (OMS) ha stabilito una soglia di sicurezza per il controllo e l'eradicazione della malattia: affinché la trasmissione del morbillo sia efficacemente contrastata, è necessario che almeno il 90% della popolazione di età superiore a un anno sia vaccinato. Questo livello di copertura è considerato fondamentale per garantire l'immunità di gregge, riducendo così il rischio di epidemie e proteggendo anche quei soggetti che, per motivi medici, non possono ricevere il vaccino. Questa prima parte dell'analisi mira quindi a valutare il grado di adesione alla vaccinazione nelle diverse regioni del mondo, individuando eventuali aree critiche dove la copertura risulta insufficiente e il rischio di diffusione della malattia rimane elevato.

Dashboard 1: Analisi dell'andamento della copertura vaccinale contro il morbillo nei continenti.

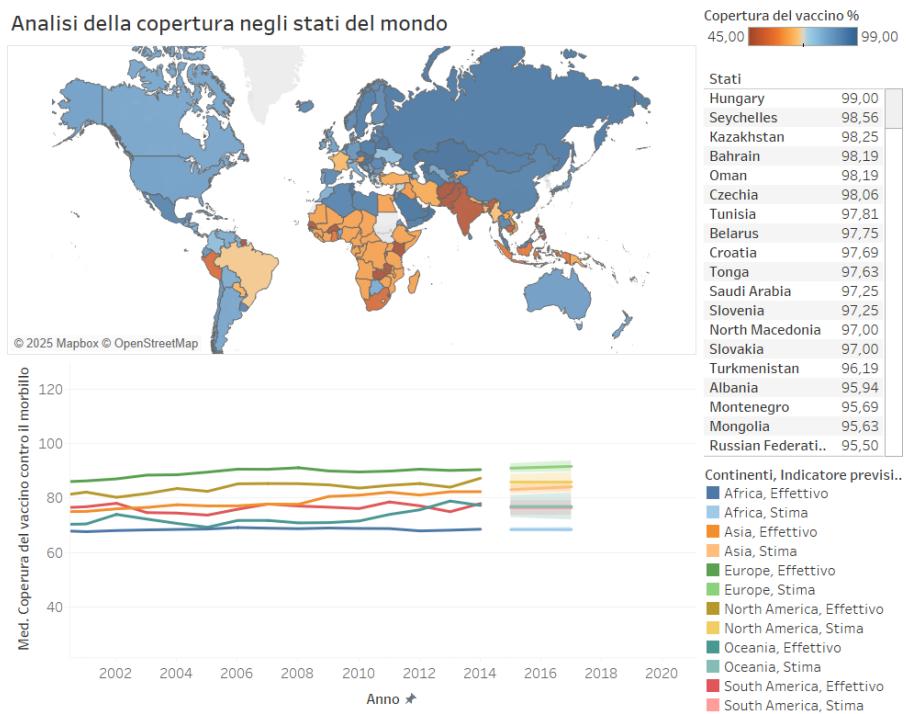


Figura 3.1

Nella **Dashboard 1** viene analizzata la copertura vaccinale contro il morbillo per ciascuno stato del mondo, calcolando la media della copertura nel periodo 2000-2015. L'obiettivo è identificare eventuali discrepanze tra le diverse aree geografiche e individuare le regioni in cui il tasso di vaccinazione risulta particolarmente critico.

Uno degli elementi più evidenti emerge dalla mappa di calore mostrata in alto a sinistra nella **Figura 3.1**, dalla quale si osserva come il continente con la copertura vaccinale più elevata sia l'Europa. Questa prima osservazione trova ulteriore conferma nel grafico in basso a sinistra, che illustra l'andamento della copertura vaccinale dal 2000 al 2015, mettendo a confronto i sei continenti analizzati. Dallo stesso grafico risulta altrettanto chiaro che l'Africa è il continente con il minor tasso di copertura vaccinale, evidenziando una persistente vulnerabilità alla diffusione della malattia. Questa osservazione suggerisce che, per contrastare efficacemente il morbillo, gli sforzi internazionali e le campagne di vaccinazione dovrebbero essere prioritariamente focalizzati sul continente afri-

cano, dove il rischio di epidemie e la mortalità associata alla malattia rimangono significativamente più alti rispetto ad altre regioni del mondo.

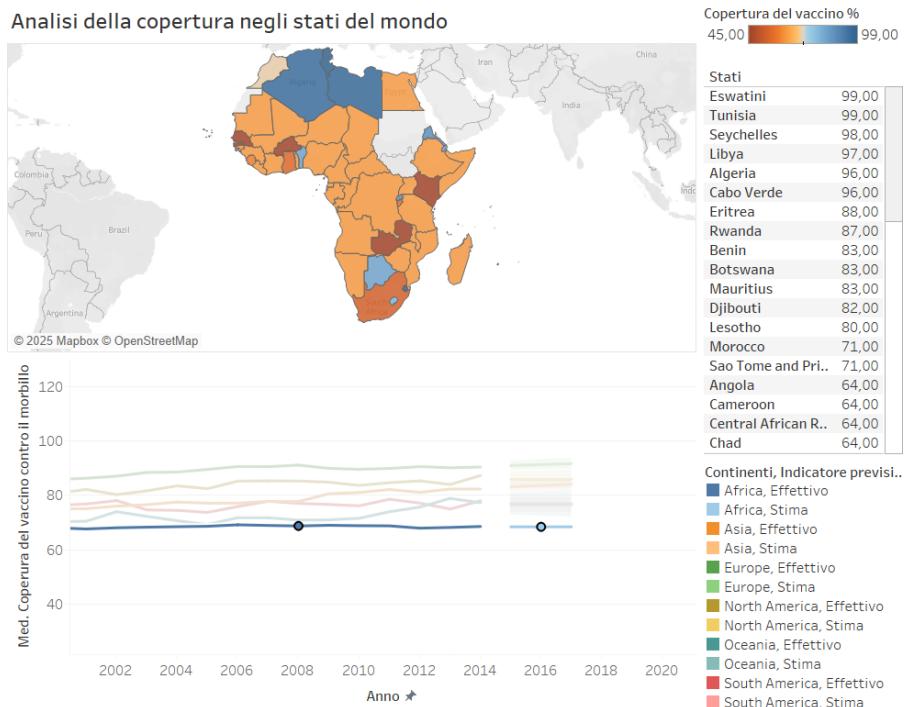


Figura 3.2

A completamento di questa analisi sui continenti, sono state considerate anche le stime per gli anni 2016-2017 relative all'andamento della copertura vaccinale. Un aspetto particolarmente interessante emerge dalla **Figura 3.2**, dove si osserva che, per l'Africa, la copertura vaccinale è stimata con un andamento costante.

Questo fenomeno può essere spiegato dal fatto che l'aumento della copertura vaccinale dipende principalmente da due fattori:

1. Lo sviluppo socioeconomico degli stati, che favorisce un accesso più ampio ai servizi sanitari, inclusa la vaccinazione.
2. Le campagne vaccinali promosse dall'OMS e da altre organizzazioni internazionali, che possono temporaneamente aumentare il tasso di copertura.

Tuttavia, poiché il primo fattore richiede un processo di crescita graduale e il secondo dipende da iniziative non regolari né sistematiche, una stima futura della copertura vaccinale non può prevedere con certezza un incremento determinato da campagne vaccinali sovrastatali, a meno che queste non vengano programmate con continuità.

Per concludere l'analisi, nella sezione di destra della dashboard, è stata realizzata una classifica degli stati con la minor copertura vaccinale media, presentandoli in ordine crescente con il relativo valore associato. Questo elenco è stato filtrato tenendo conto dei risultati precedentemente evidenziati sull'Africa, permettendo di individuare in modo chiaro le nazioni che richiedono un'attenzione prioritaria per il miglioramento delle strategie di immunizzazione.

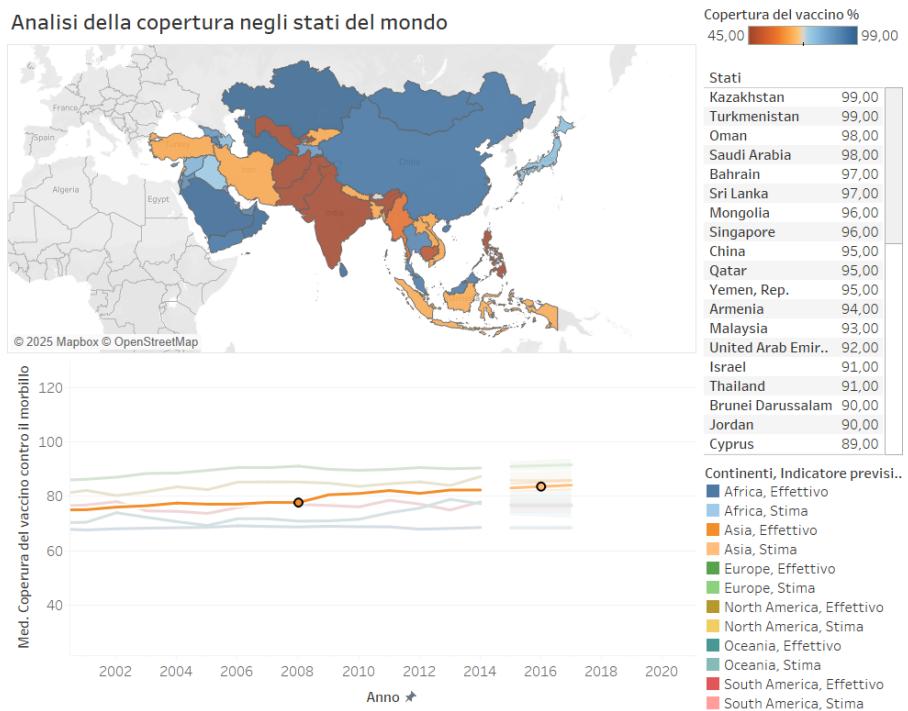


Figura 3.3

A differenza dell'Africa, che mostra un andamento della copertura vaccinale relativamente stabile, il continente con la maggiore crescita percentuale nella vaccinazione contro il morbillo è l'Asia, come evidenziato nel diagramma in basso a sinistra della **Figura 3.3**.

L'Asia si distingue sia per l'aumento effettivo della copertura vaccinale registrato tra il 2000 e il 2015, sia per le stime di crescita per gli anni successivi. Questo trend positivo è indicativo di un progressivo sviluppo socioeconomico nel continente, che ha contribuito a migliorare l'accesso ai servizi sanitari e a rafforzare le politiche di immunizzazione. Questa tendenza suggerisce che, nel lungo periodo, l'Asia potrebbe raggiungere livelli di copertura vaccinale comparabili a quelli dei paesi più sviluppati, contribuendo alla riduzione globale del morbillo e delle sue complicanze.

Dashboard 2: Analisi dell'andamento della copertura del vaccino contro il morbillo nelle regioni.

Copertura del vaccino nelle regioni

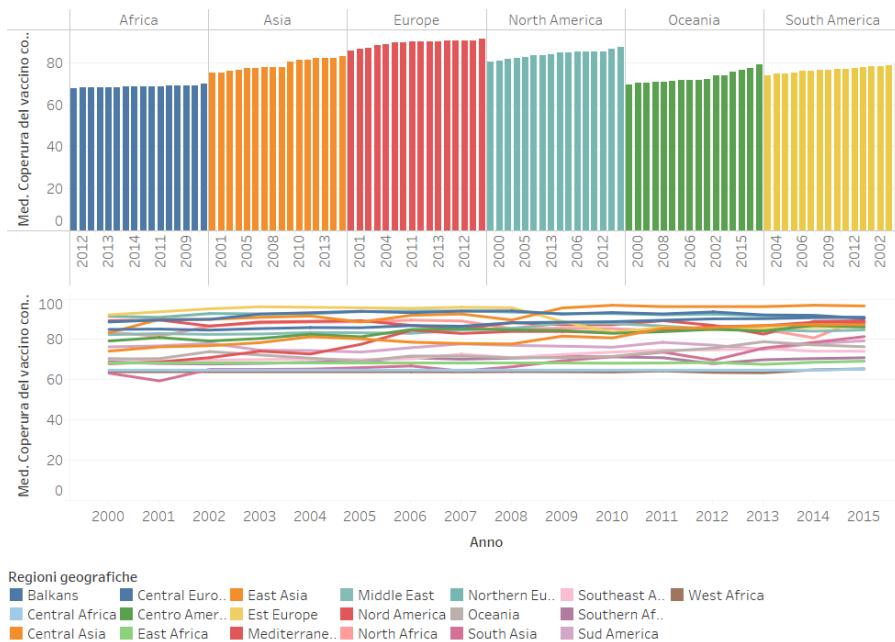


Figura 3.4

Dopo aver esaminato l'andamento della copertura vaccinale contro il morbillo a livello continentale nella **Dashboard 1**, è fondamentale approfondire l'analisi distinguendo tra le diverse regioni all'interno di ciascun continente. Questo permette di individuare con maggiore precisione le aree geografiche che presentano criticità nella copertura vaccinale e di comprendere meglio le dinamiche regionali che influenzano la diffusione della malattia.

La **Dashboard 2** è progettata proprio per questo scopo. Nella parte superiore, essa mostra l'andamento della copertura vaccinale per ciascun continente nel periodo compreso tra il 2000 e il 2015, consentendo una visione d'insieme dei trend globali.

Un elemento chiave di questa dashboard è la possibilità di selezionare un singolo continente per ottenere un'analisi più dettagliata. Una volta effettuata la selezione, nella parte inferiore della **Figura 3.4** vengono evidenziati gli andamenti delle singole regioni che compongono il continente scelto. Questa suddivisione permette di:

- Identificare differenze significative tra le regioni all'interno dello stesso continente.
- Osservare le tendenze specifiche di ciascuna area geografica, evidenziando eventuali miglioramenti o stagnazioni nella copertura vaccinale.
- Individuare le regioni prioritarie per future campagne di vaccinazione e interventi sanitari mirati.

Questo livello di dettaglio offre quindi una visione più approfondita del fenomeno, permettendo di andare oltre l'analisi generale a livello continentale e di comprendere le dinamiche locali che influenzano l'accesso e l'adesione ai programmi vaccinali.

Copertura del vaccino nelle regioni



Figura 3.5

Seguendo le analisi precedenti sull'Africa e integrando i dati sugli andamenti delle regioni africane, emerge con chiarezza una forte variabilità regionale nella copertura vaccinale contro il morbillo.

Come evidenziato nella parte inferiore della **Figura 3.5**, dopo aver selezionato il continente africano nella dashboard, risulta immediatamente evidente che la regione del Nord Africa presenta una copertura vaccinale significativamente

superiore rispetto alle altre aree del continente. In particolare, il tasso di copertura in questa regione è mediamente superiore di circa venti punti percentuali rispetto alle regioni dell'Africa Centrale, Orientale, Occidentale e Meridionale.

Questa osservazione assume un'importanza cruciale dal punto di vista decisionale e strategico per la pianificazione di campagne vaccinali mirate. Infatti, dato che il Nord Africa ha già raggiunto una copertura in linea con gli standard raccomandati dall'OMS, gli sforzi per migliorare l'immunizzazione dovrebbero essere prioritariamente concentrati sulle regioni con tassi di copertura inferiori.

Pertanto, una campagna vaccinale efficace in Africa dovrebbe focalizzarsi sulle seguenti aree: Africa Centrale, Africa Orientale, Africa Occidentale, Africa Meridionale

Queste regioni, caratterizzate da livelli di vaccinazione ancora insufficienti, rappresentano i territori più vulnerabili alla diffusione del morbillo e, di conseguenza, quelli in cui interventi mirati potrebbero avere un impatto significativo nella riduzione della mortalità infantile legata alla malattia.

Copertura del vaccino nelle regioni

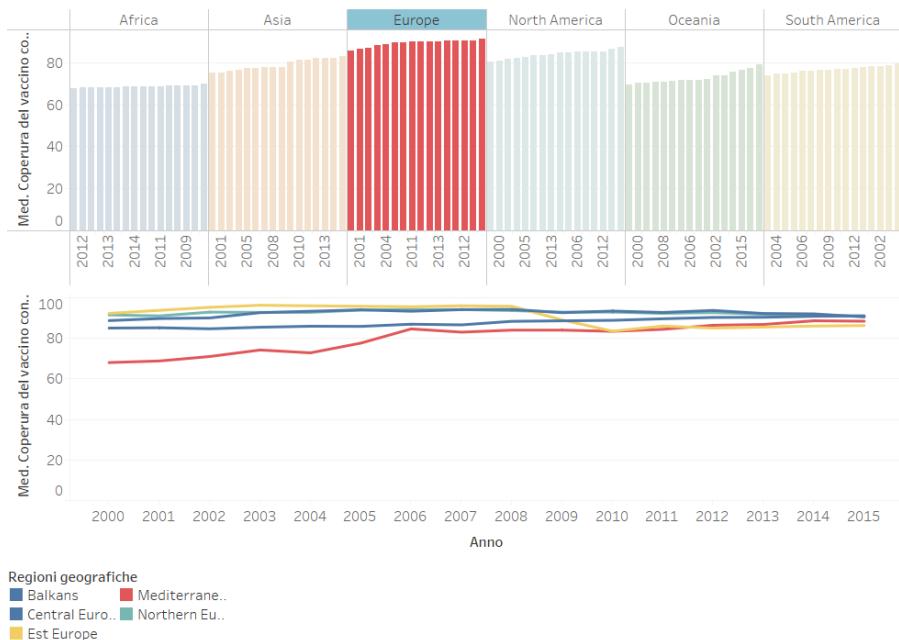


Figura 3.6

A differenza dell'Africa, in Europa i livelli di copertura vaccinale contro il morbillo hanno già raggiunto gli standard di sicurezza raccomandati dall'OMS. Tuttavia, l'analisi dei dati rivela la presenza di due regioni con andamenti ano-

mali, che meritano un approfondimento specifico, come evidenziato nella **Figura 3.6**.

Nei primi anni 2000, la regione Mediterranea mostrava un livello di copertura relativamente basso rispetto agli standard europei. Questo fenomeno è riconducibile principalmente all'andamento della vaccinazione in paesi come Malta e Portogallo, che inizialmente registravano valori inferiori alla media europea. Tuttavia, negli anni successivi, entrambi gli stati hanno implementato strategie di recupero e si sono progressivamente allineati ai livelli di copertura vaccinale europei, garantendo una protezione adeguata alla popolazione.

Un'altra anomalia significativa è osservabile nell'Europa dell'Est, che per diversi anni ha mantenuto una copertura vaccinale mediamente elevata, ma ha poi subito un crollo improvviso nel 2009. Questo calo è stato direttamente influenzato dalla crisi economica che ha colpito alcuni paesi dell'area, con ripercussioni sulle politiche sanitarie e sulla capacità di mantenere elevati livelli di immunizzazione nella popolazione.

L'analisi di questo fenomeno sarà approfondita nelle sezioni successive, dove verranno esaminati più nel dettaglio gli effetti della crisi economica sulla copertura vaccinale e sulle possibili strategie per evitare che situazioni simili si ripetano in futuro.

Dashboard 3: Analisi dell'andamento della copertura del vaccino contro il morbillo nei singoli stati

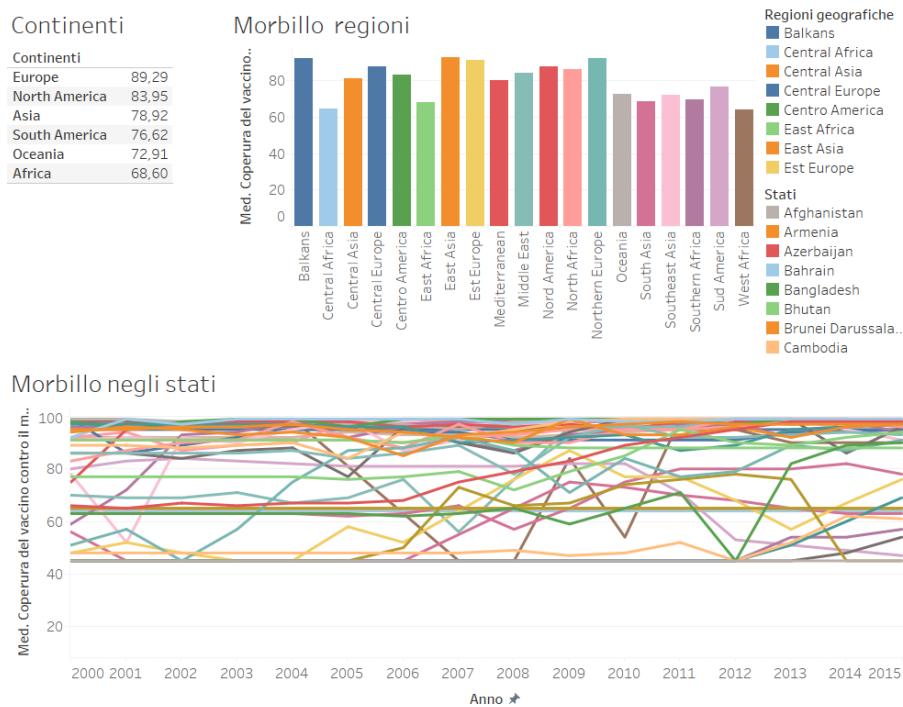


Figura 3.7

Dopo aver analizzato l'andamento delle regioni geografiche nella **Dashboard 2**, è possibile affinare ulteriormente il livello di dettaglio dell'analisi, passando da una visione regionale a una più specifica per singolo stato.

Nella **Figura 3.7**, la parte superiore mostra nuovamente la suddivisione dei dati per continenti e regioni, come già visto in precedenza, con le medie di copertura vaccinale calcolate per il periodo 2000-2015. Questo consente di confermare i trend generali già individuati nelle fasi precedenti dell'analisi.

Tuttavia, l'elemento più significativo di questa dashboard si trova nella parte inferiore della figura, dove viene rappresentato l'andamento della copertura vaccinale per ogni singolo stato nel medesimo arco temporale. Questo approfondimento permette di individuare con precisione quali stati hanno mantenuto una copertura costante, quali hanno registrato miglioramenti significativi e quali, invece, presentano criticità persistenti.

Grazie alla struttura interattiva della **Dashboard 3**, è possibile raffinare l'analisi partendo dal livello continentale e scendendo progressivamente fino al singolo stato, passando per la regione geografica di appartenenza. Questo ap-

proccio consente di identificare le aree che necessitano di interventi mirati per migliorare la copertura vaccinale e prevenire la diffusione del morbillo.

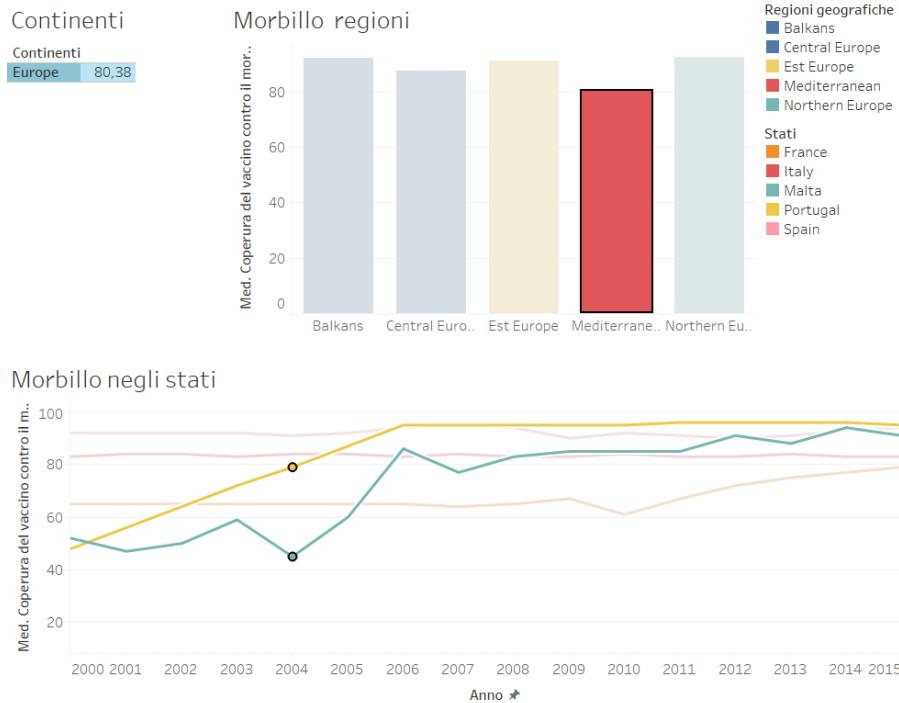


Figura 3.8

Come già anticipato, un'analisi più dettagliata della zona mediterranea evidenzia alcune anomalie interessanti. In particolare, si può osservare che, nei primi anni 2000, alcuni stati come Malta e Portogallo si trovavano significativamente al di sotto della media europea per quanto riguarda la copertura vaccinale contro il morbillo.

Questa discrepanza, chiaramente visibile nella **Figura 3.8**, rappresentava un potenziale fattore di rischio per la diffusione della malattia in questi paesi. Tuttavia, lo stesso grafico mostra come, negli anni successivi, entrambi gli stati abbiano progressivamente incrementato la propria copertura vaccinale, allineandosi gradualmente agli standard europei.

Grazie a interventi mirati e strategie sanitarie efficaci, Malta e Portogallo sono riusciti a raggiungere il 90% della popolazione vaccinata, il valore di sicurezza raccomandato dall'OMS per garantire un'adeguata protezione contro il morbillo. Questo esempio dimostra come politiche sanitarie mirate e campagne di sensibilizzazione possano colmare il divario e portare gli stati con copertura insufficiente ai livelli ottimali di immunizzazione.

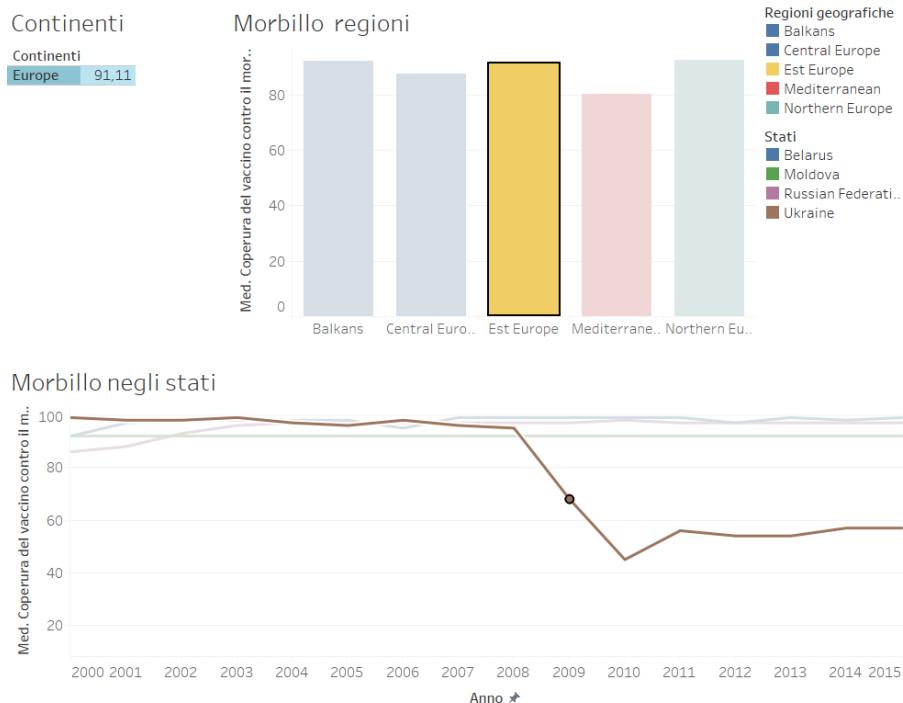


Figura 3.9

Un'analisi specifica dell'Europa dell'Est evidenzia un caso di particolare interesse: l'Ucraina. Osservando il grafico in **Figura 3.9**, si nota come fino al 2008 la copertura vaccinale contro il morbo di pozzillo fosse molto elevata, per poi subire un drastico crollo negli anni successivi.

Per comprendere le cause di questa diminuzione, è utile confrontare l'andamento della copertura vaccinale con il contesto storico e politico del paese. Infatti, il calo coincide con l'intensificarsi delle tensioni con la Russia, che hanno portato a una crescente instabilità politica ed economica. Questa situazione ha probabilmente influito negativamente sulle politiche sanitarie e sulla gestione dei programmi di vaccinazione, determinando una riduzione della copertura.

Approfondendo ulteriormente il grafico, si osserva che prima del 2008, tutti gli stati della regione presentavano livelli molto alti di copertura vaccinale. Sebbene non sia direttamente visibile nel grafico, questi paesi condividono una caratteristica comune: appartenevano all'ex Unione Sovietica. Questo suggerisce che, almeno fino ai primi anni 2000, il sistema sanitario ereditato dal periodo sovietico garantisse una copertura vaccinale omogenea e relativamente alta in tutta la regione.

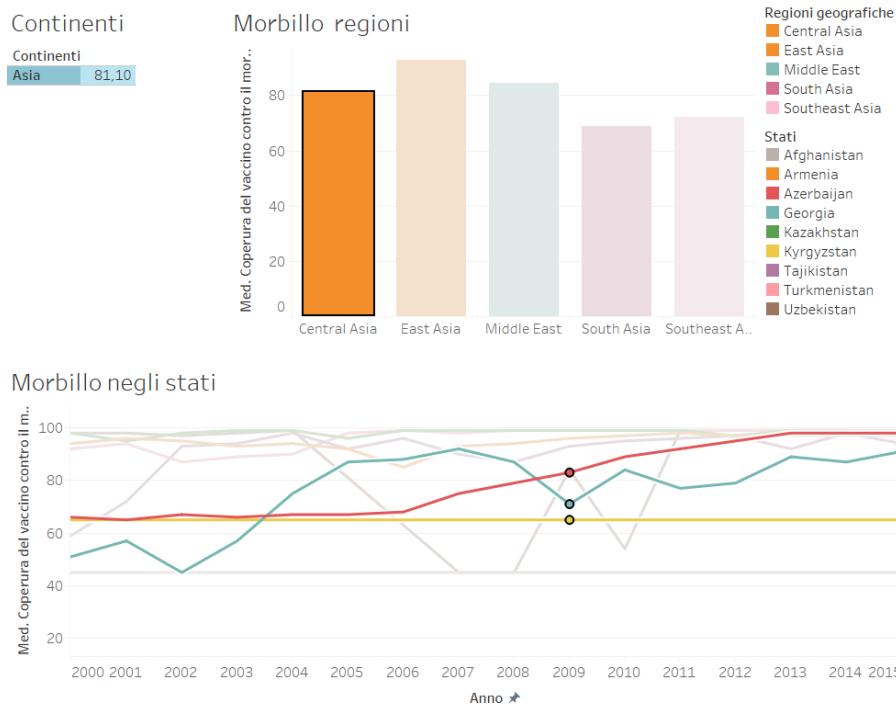


Figura 3.10

Unendo queste informazioni con un'analisi storica più ampia, è possibile individuare schemi simili in altri stati dell'ex Unione Sovietica che hanno affrontato conflitti armati.

Il grafico in **Figura 3.10** evidenzia l'andamento della copertura vaccinale in tre paesi specifici: Georgia, Azerbaigian e Kirghizistan. In questi casi, i conflitti si sono verificati prima del 2000, quindi nel grafico non si osserva un crollo improvviso della copertura, ma piuttosto la sua evoluzione successiva.

Un aspetto interessante emerge analizzando il caso del Kirghizistan. A differenza degli altri due paesi, la sua copertura vaccinale mostra un andamento costante e relativamente basso. Questo comportamento può essere collegato alla presenza di conflitti interni ancora irrisolti, che continuano a influenzare negativamente il sistema sanitario nazionale. L'instabilità politica e le difficoltà economiche impediscono una crescita significativa della copertura vaccinale, rendendo il paese più vulnerabile alla diffusione del morbo.

Dashboard 4: Analisi della relazione tra copertura del vaccino contro il morbillo e morti nei bambini

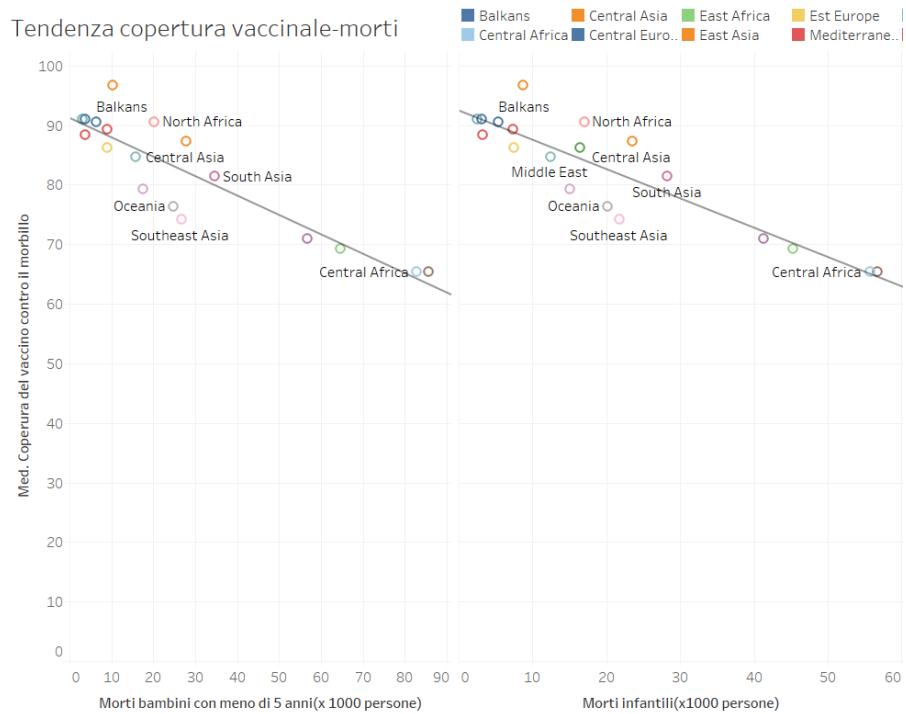


Figura 3.11

Per completare l'analisi sulla copertura vaccinale contro il morbillo, è essenziale esaminare la relazione tra tassi di vaccinazione e mortalità infantile.

Nella **Dashboard 4**, i dati vengono rappresentati attraverso due grafici distinti: A sinistra: il numero di morti nei bambini sotto i 5 anni ogni mille nati. A destra: il numero di morti nei bambini sotto l'anno di età.

Osservando le tendenze in **Figura 3.11**, emerge un legame evidente tra copertura vaccinale e mortalità infantile. In particolare, la linea di tendenza mostra una pendenza negativa, indicando che aumentando la copertura vaccinale, la mortalità infantile diminuisce.

Analizzando più in dettaglio le scale di misura, si può dedurre che questa relazione è valida sia per i bambini sotto i 5 anni sia per quelli sotto l'anno di età, confermando l'importanza della vaccinazione nel ridurre il rischio di morte legato a malattie prevenibili come il morbillo. Questi dati sottolineano ancora una volta l'urgenza di implementare strategie sanitarie efficaci, soprattutto nelle regioni con basse coperture vaccinali, per ridurre la mortalità infantile a livello globale.

3.6 Analisi sul tasso di mortalità nei bambini con meno di 5 anni e ricerca di possibili cause

Dopo aver individuato una relazione tra la copertura vaccinale contro il morbillo e il tasso di mortalità nei bambini di età inferiore ai cinque anni, sorge spontanea la domanda se vi siano altri fattori che contribuiscono a questi decessi. L'analisi si estenderà dunque a una serie di variabili potenzialmente rilevanti, suddivise in tre macrocategorie: fattori geografici, economici e sanitari.

In particolare, per quanto riguarda gli aspetti sanitari, si prenderanno in esame le coperture vaccinali relative a tre diverse malattie, escludendo il morbillo, già oggetto di studio. Questo permetterà di valutare se e in che misura la protezione immunitaria fornita da altri vaccini possa incidere sulla riduzione della mortalità infantile. L'obiettivo è fornire un quadro più completo delle cause che influenzano la sopravvivenza nei primi anni di vita e identificare eventuali strategie di intervento mirate.

Dashboard 5: Analisi della distribuzione geografica delle morti nei bambini con meno di 5 anni

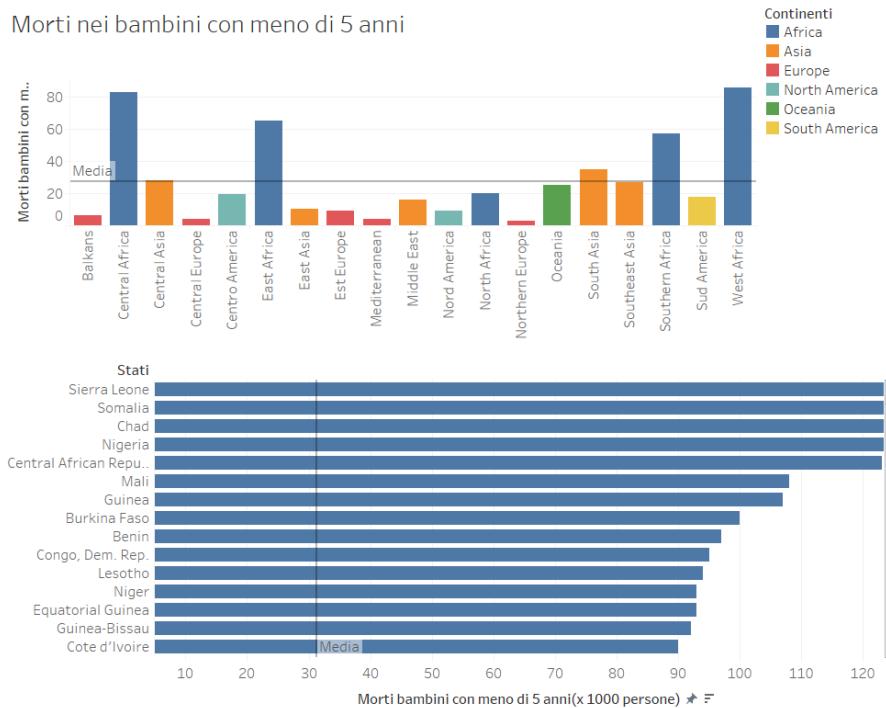


Figura 3.12

Nella **Dashboard 5** viene presentata un'analisi dettagliata della mortalità infantile nei bambini sotto i cinque anni, suddivisa dapprima per regioni e successivamente per singoli stati. Ogni continente è identificato da un colore specifico, facilitando l'interpretazione visiva dei dati.

Dal grafico principale emerge chiaramente come le regioni più colpite siano prevalentemente quelle africane. Tuttavia, come già osservato in precedenza, il Nord Africa si distingue dalle altre aree del continente, registrando un numero di decessi relativamente inferiore. Questo suggerisce la presenza di fattori specifici, come migliori condizioni sanitarie ed economiche, che potrebbero contribuire a questa differenza all'interno dello stesso continente.

Nel medesimo grafico è stata calcolata la media dei decessi per regione, ma tale valore non è stato incluso direttamente nella visualizzazione, poiché non apportava un'informazione aggiuntiva significativa. Infatti, l'andamento dei dati mostra chiaramente che poche regioni si avvicinano al valore medio. Al contrario, si osservano forti disparità, con alcune regioni che registrano picchi

molto elevati di mortalità infantile e altre che presentano valori sensibilmente più bassi.

La **Figura 3.12** fornisce un’ulteriore analisi dettagliata, ordinando in modo decrescente gli stati in base al numero di decessi registrati. Grazie alla codifica cromatica che associa ogni stato al proprio continente, risulta immediatamente evidente che i primi venti paesi con il più alto numero di morti infantili appartengono esclusivamente al continente africano. Nessun altro continente compare nelle posizioni più alte della graduatoria, sottolineando così il forte impatto della mortalità infantile in questa parte del mondo rispetto alle altre regioni globali.

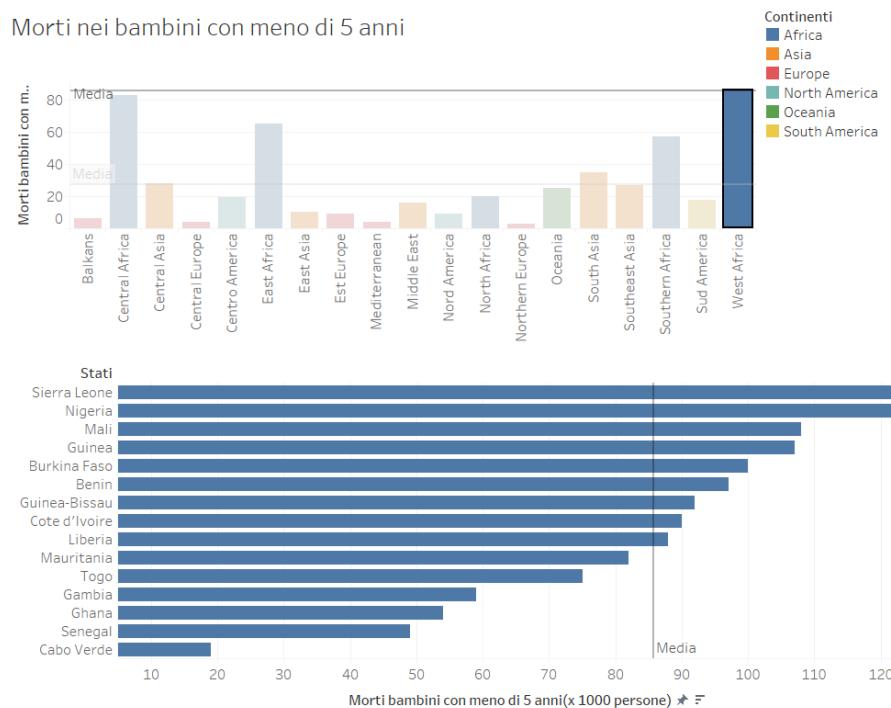


Figura 3.13

Nella **Figura 3.13** è stata applicata un’analisi per regione, evidenziando come all’interno dell’Africa occidentale siano presenti significative differenze nei tassi di mortalità infantile tra i vari stati. Un esempio particolarmente virtuoso è rappresentato da Capo Verde, che registra valori sensibilmente più bassi rispetto alla media della regione.

Dallo stesso grafico emerge, tuttavia, che all’interno della stessa area geografica si trovano anche paesi con tassi di mortalità estremamente elevati, come la Sierra Leone e la Nigeria. Questa forte disomogeneità suggerisce che fattori specifici, oltre al contesto regionale, possano influenzare la sopravvivenza infantile.

Una possibile spiegazione per il caso di Capo Verde è il ruolo del turismo: essendo la principale destinazione turistica della regione, è plausibile che una parte degli introiti derivanti da questo settore contribuisca al miglioramento delle condizioni sanitarie e socioeconomiche, riducendo così il tasso di mortalità infantile.

Un'osservazione simile può essere fatta analizzando la **Figura 3.14**, che si concentra sulla situazione dell'Africa meridionale. Anche in questa regione si riscontrano variazioni significative nei tassi di mortalità, con alcuni stati che mostrano valori più contenuti rispetto ad altri. Questo suggerisce che fattori economici, come la presenza di settori produttivi e turistici sviluppati, possano incidere positivamente sulle condizioni di vita e sulla salute della popolazione infantile.

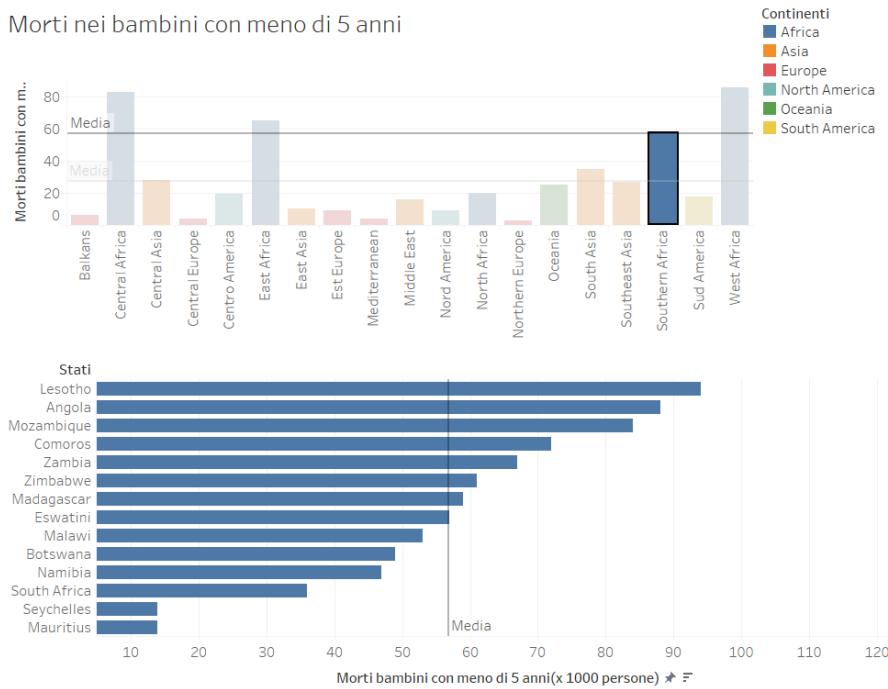


Figura 3.14

Dashboard 6: Analisi della correlazione tra morti nei bambini con meno di 5 anni e GDP per capita

Nella **Dashboard 6** è stata effettuata un'analisi delle diverse regioni del mondo con l'obiettivo di individuare una possibile relazione tra la mortalità infantile media (nei bambini sotto i cinque anni) e il prodotto interno lordo pro capite (GDP per capita).

Questa analisi consente di valutare se e in che misura il livello di ricchezza economica di una regione influisce sulla sopravvivenza infantile. In particolare, confrontando i dati delle diverse aree geografiche, è possibile osservare eventuali tendenze che suggeriscono un legame tra sviluppo economico e condizioni sanitarie, individuando regioni che si discostano da questa relazione generale.

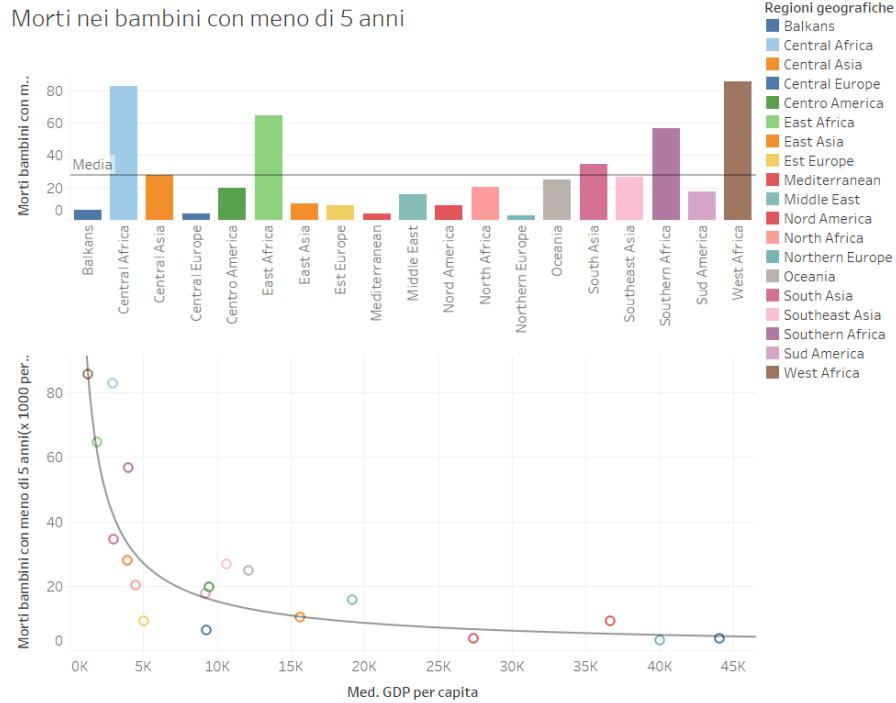


Figura 3.15

Nella **Figura 3.15** è stata individuata una relazione di tipo potenza tra il GDP pro capite e la mortalità infantile media nelle diverse regioni analizzate. Questo andamento indica che, con l'aumentare del reddito medio per abitante, il numero di decessi nei bambini sotto i cinque anni diminuisce progressivamente, fino a stabilizzarsi attorno a un valore costante.

In particolare, l'analisi mostra che al di sotto di una certa soglia economica, un aumento del GDP pro-capite è associato a una diminuzione significativa della mortalità infantile. Tuttavia, superata una determinata soglia, in questo caso un valore inferiore a 20 morti ogni 1.000, ulteriori incrementi del reddito non portano a una riduzione significativa delle morti infantili, suggerendo un effetto di saturazione. Questo implica che, oltre un certo livello di sviluppo economico, altri fattori (come il miglioramento delle strutture sanitarie e l'accesso ai servizi di base) diventano determinanti nel mantenere bassi i tassi di mortalità, piuttosto che la sola crescita economica.

L'attendibilità di questo modello è rafforzata dal fatto che tutte le regioni analizzate si distribuiscono in modo omogeneo lungo la curva di tendenza, senza evidenziare valori anomali o punti di forte deviazione. Questo suggerisce che la relazione individuata è solida e generalizzabile, offrendo un'importante chiave di lettura sul legame tra sviluppo economico e salute infantile a livello globale.

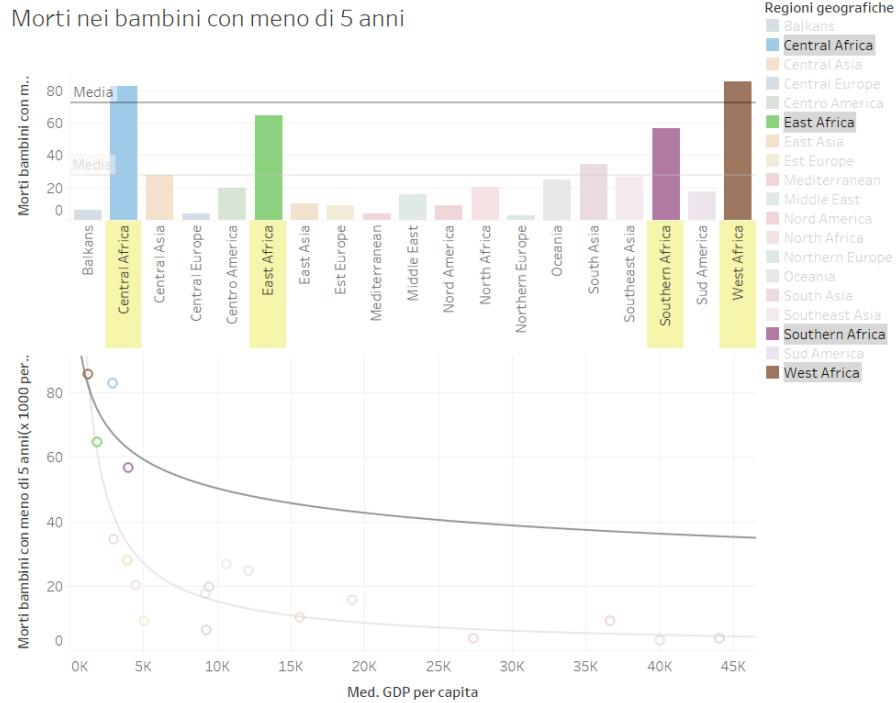


Figura 3.16

Ancora una volta, risulta particolarmente interessante analizzare il comportamento di alcune regioni africane in relazione alla mortalità infantile e al GDP pro capite. Nella **Figura 3.16**, si osserva che queste regioni si collocano prevalentemente nella parte alta della curva, caratterizzata da un elevato tasso di mortalità e da un basso reddito pro capite.

Secondo questa analisi, un aumento del GDP pro capite in queste aree potrebbe portare a una significativa riduzione della mortalità infantile, come suggerito dall'andamento della curva di tendenza individuata in precedenza. Questo rafforza l'ipotesi che il miglioramento delle condizioni economiche possa avere un impatto diretto sulla sopravvivenza infantile, specialmente nei contesti in cui il reddito medio è ancora molto basso.

Tale dinamica diventa ancora più evidente quando si introduce nel confronto la regione del Nord Africa, come mostrato nella **Figura 3.17**. Quest'area presenta un GDP pro capite superiore rispetto al resto del continente e, coeren-

temente con il modello osservato, registra un tasso di mortalità inferiore. Questo confronto supporta ulteriormente l'idea che il progresso economico possa essere un fattore determinante nella riduzione della mortalità infantile, almeno fino a una certa soglia oltre la quale altri fattori assumono un ruolo predominante.

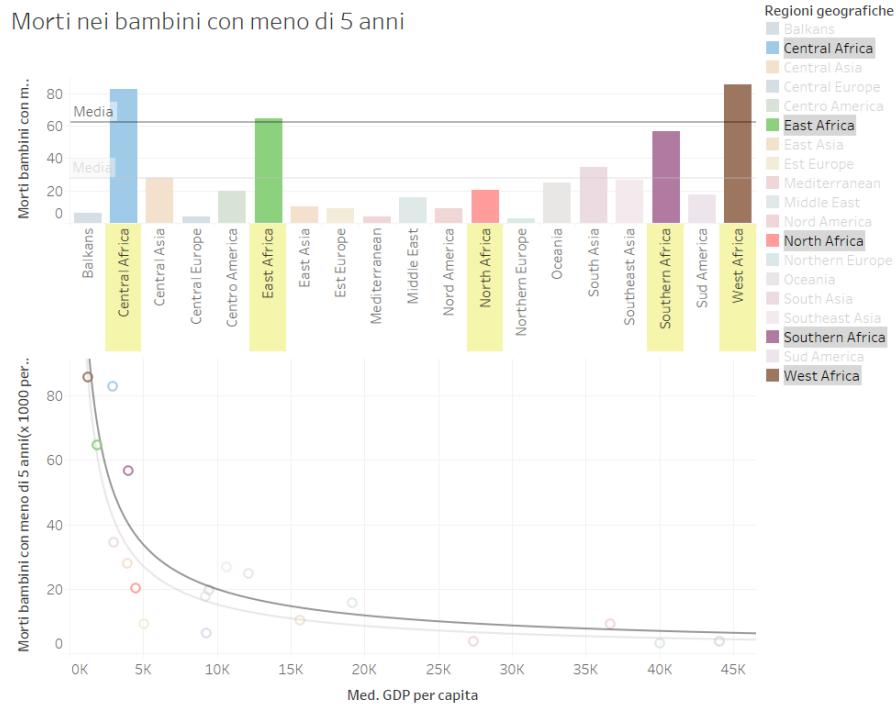


Figura 3.17

Dashboard 7: Analisi della correlazione tra le coperture vaccinali e le morti nei bambini con meno di 5 anni

Per concludere l'analisi sulla mortalità infantile nei bambini sotto i cinque anni, nella **Dashboard 7** è stata esaminata la relazione tra il tasso di mortalità e la copertura vaccinale per tre malattie specifiche: poliomielite, epatite B e difterite.

L'obiettivo di questa analisi è valutare se esiste un'associazione tra l'estensione dei programmi vaccinali e la riduzione della mortalità infantile nelle diverse regioni del mondo. A tal fine, per ciascuna delle tre malattie considerate, è stato calcolato e inserito nel grafico il valore medio di copertura vaccinale per ogni regione nel periodo compreso tra il 2000 e il 2015.

Questa rappresentazione permette di osservare eventuali tendenze e discrepanze tra le diverse aree geografiche, evidenziando quali regioni abbiano beneficiato maggiormente da un'ampia diffusione dei vaccini e quali, invece, presentino

ancora livelli di copertura insufficienti. L'analisi fornisce così un quadro più ampio dell'impatto dei programmi di immunizzazione sulla sopravvivenza infantile, contribuendo a comprendere meglio il ruolo della sanità pubblica nella riduzione della mortalità nei primi anni di vita.

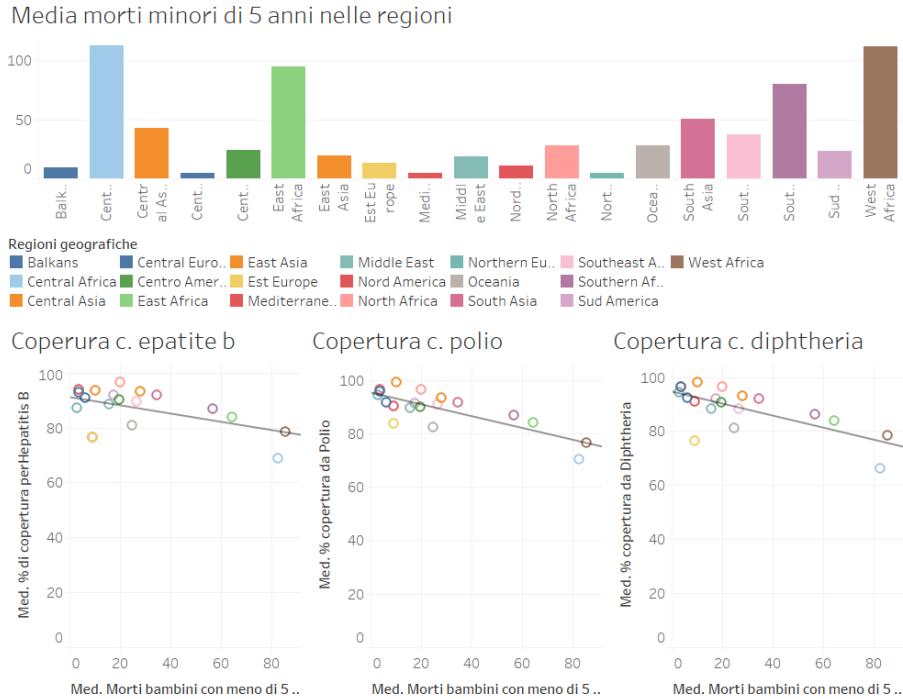


Figura 3.18

Analogamente a quanto osservato per il morbillo, anche in questo caso è possibile intuire una correlazione tra la copertura vaccinale e la mortalità infantile nei bambini sotto i cinque anni. Analizzando in modo più dettagliato i dati presentati nei grafici della **Figura 3.18**, si nota che questa relazione risulta particolarmente forte nei casi di difterite e poliomielite. Questo è evidente dalla maggiore pendenza della linea di tendenza, che suggerisce un impatto più marcato della copertura vaccinale nella riduzione della mortalità infantile per queste due malattie.

Dai medesimi grafici emerge anche un altro dato rilevante: le regioni africane continuano a registrare i valori peggiori in termini di mortalità infantile e copertura vaccinale. Tuttavia, queste regioni seguono comunque l'andamento generale evidenziato nel resto del mondo, confermando che la vaccinazione è un fattore chiave nella riduzione dei decessi nei bambini sotto i cinque anni.

Questa osservazione porta a una conclusione più ampia: l'elevata mortalità infantile in Africa non è determinata da un singolo fattore, ma piuttosto dal-

l'interazione di molteplici elementi, tra cui bassa copertura vaccinale, condizioni sanitarie precarie, limitato accesso alle cure mediche e fattori economici e sociali sfavorevoli. Questo spiega perché, nonostante la tendenza globale indichi una correlazione tra vaccinazione e riduzione della mortalità, le regioni africane continuano a registrare tassi di mortalità significativamente più alti rispetto al resto del mondo.

3.7 Analisi geografica sulla malnutrizione nei bambini tra i 5 ed i 19 anni e ricerca di possibili correlazioni

Dashboard 8: Analisi geografica della malnutrizione nei continenti, nelle regioni e negli stati

Dopo aver esaminato il tasso di mortalità nei bambini sotto i 5 anni in relazione a fattori geografici, economici e sanitari, è fondamentale analizzare un'altra problematica che colpisce milioni di bambini nel mondo: la malnutrizione. Nella **Dashboard 8**, viene approfondito il tasso di malnutrizione tra i bambini di età compresa tra i 5 e i 19 anni nelle diverse regioni del mondo. Nella parte superiore del grafico, i dati sono suddivisi per continente e regione, mentre nella parte inferiore è riportato il tasso di malnutrizione per ciascun stato, consentendo un'analisi più dettagliata della distribuzione del fenomeno.

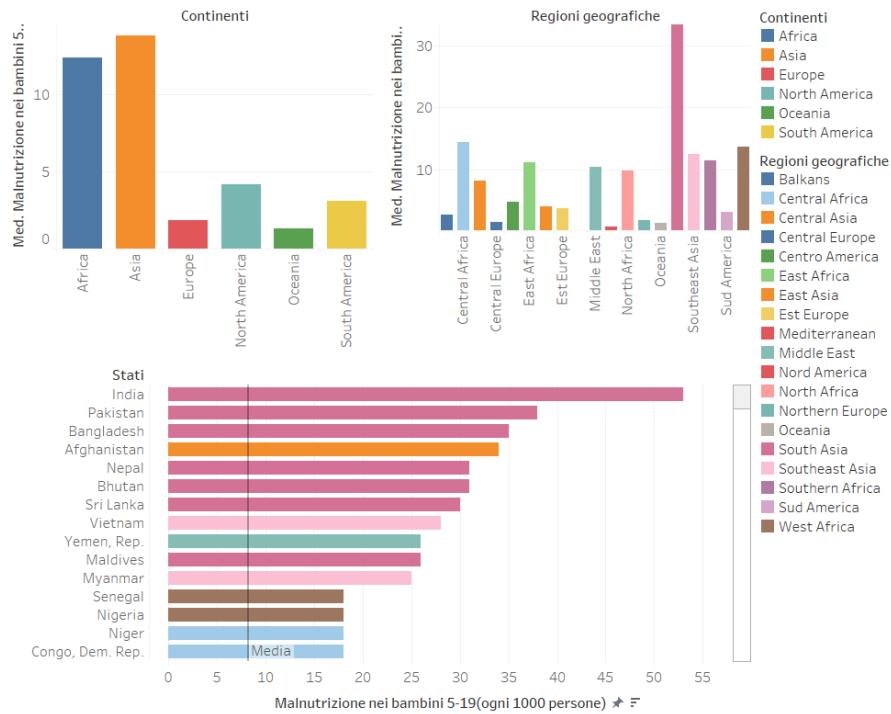


Figura 3.19

Osservando il grafico della **Figura 3.19**, nella parte in alto a sinistra, emerge un dato sorprendente: a differenza di altre problematiche analizzate in precedenza, il tasso di malnutrizione più elevato non si registra in Africa, bensì in Asia. Approfondendo l'analisi con il grafico in alto a destra, spicca in particolare la regione del Sud Asia, che presenta i valori più alti, mentre le altre regioni asiatiche mostrano tassi di malnutrizione più in linea con quelli africani. Inoltre, la parte inferiore della dashboard evidenzia come i tre stati con il più alto tasso di malnutrizione al mondo siano India, Pakistan e Bangladesh. Complessivamente, tra i primi dieci paesi con i livelli più elevati di malnutrizione, figurano esclusivamente stati asiatici.

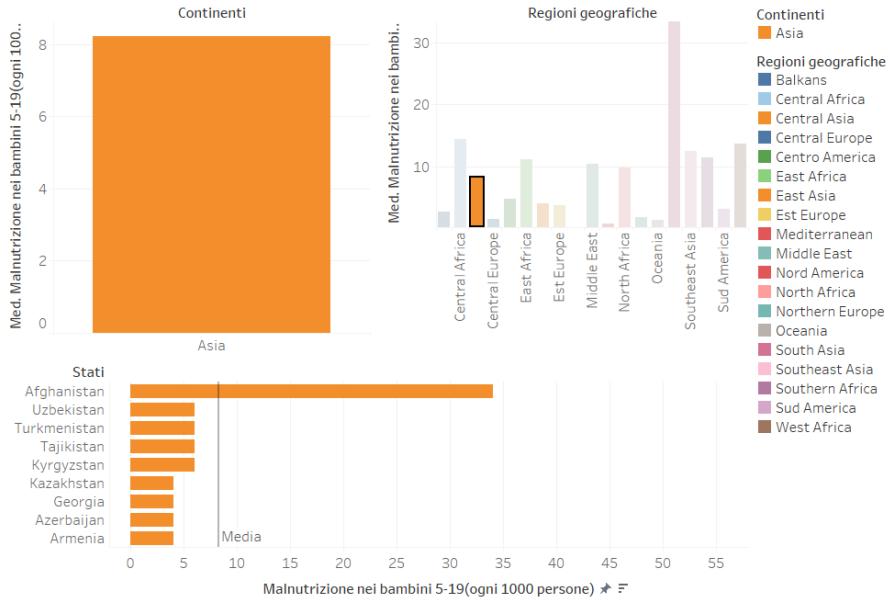


Figura 3.20

La **Figura 3.20** evidenzia un'anomalia nell'elenco degli stati. Sebbene il tasso di malnutrizione nella regione dell'Asia occidentale sia generalmente più basso rispetto alla media continentale, un'eccezione significativa è rappresentata dall'Afghanistan, che si distingue per valori nettamente superiori rispetto agli altri paesi della regione. Approfondendo l'analisi, è evidente che questo dato anomalo può essere ricondotto alle conseguenze del conflitto tra Afghanistan e Stati Uniti tra il 2001 e il 2015, che ha avuto un impatto devastante sulle condizioni economiche e sanitarie della popolazione.

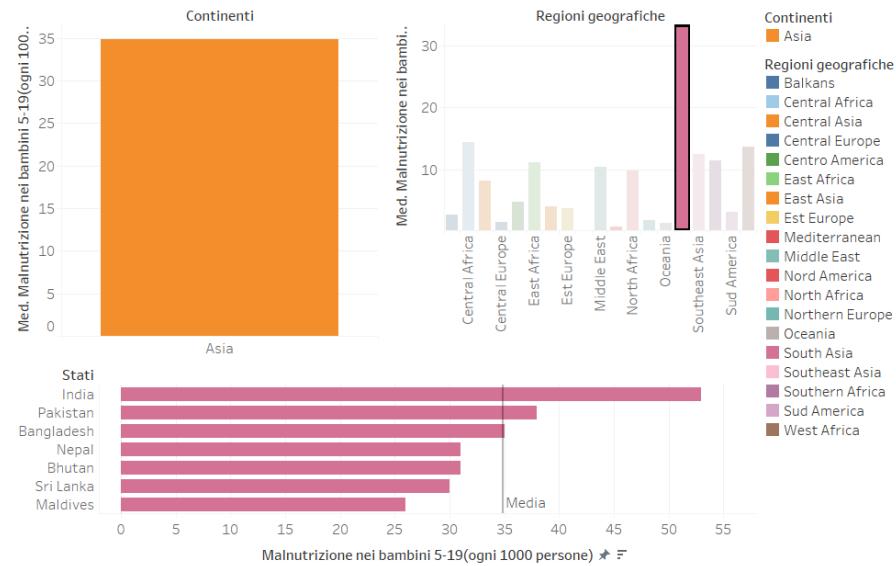


Figura 3.21

Nella **Figura 3.21**, viene approfondita la situazione nel Sud Asia, dove si osserva una maggiore uniformità nei tassi di malnutrizione rispetto ad altre regioni. Tuttavia, l'India continua a distinguersi con valori superiori rispetto agli altri stati dell'area. In questo caso, l'assenza di conflitti estesi nella regione suggerisce che il principale fattore determinante sia la densità della popolazione indiana, significativamente più alta rispetto agli altri paesi vicini. Con oltre un miliardo di abitanti, l'India affronta sfide strutturali che influenzano direttamente la distribuzione delle risorse alimentari e l'accesso a un'alimentazione adeguata per tutti i suoi cittadini.

Capitolo 4

Power BI

4.1 Introduzione

Power BI è una piattaforma di Business Intelligence e analytics sviluppata da Microsoft che permette di analizzare, visualizzare, e condividere dati aziendali in modo efficace. Power BI offre una vasta gamma di strumenti software e app per creare report interattivi e dashboard. Questo offre agli utenti una visione chiara delle tendenze e delle informazioni chiave, grazie alle quali possono individuare problemi e prendere decisioni formate.

4.1.1 ETL: Extract, Transform e Load

Nella prima fase del processo di ETL i dati sono stati scaricati dalla sorgente e importati in formato CSV nell'ambiente Power BI. Terminato il processo di estrazione sono stati utilizzati strumenti come l'editor di query *Power Query* per la seconda fase di trasformazione dei dati. Tra le principali modifiche che si è reso necessario apportare vi sono:

- **Conversione dei tipi di dati:** durante la fase di caricamento è emersa una problematica nella colonna *Sleep Duration* che convertiva autonomamente i valori da float a interi. A tal proposito abbiamo creato una nuova colonna di tipo numerico *New Sleep Duration* utilizzando lo script presente in *Figura 4.1, Table.AddColumn*, nella quale sono stati convertiti i valori errati.
- **Creazione di nuovi attributi:** ai fini dell'analisi sono state create nuove colonne a partire da altre già esistenti. La prima, *Category Stress Level*, è stata generata dalla discretizzazione della colonna *Stress Level*, suddividendo i valori in tre categorie: *Alto* per valori superiori a 6, *Medio* per valori compresi tra 5 e 6 e, infine, *Basso* per valori inferiori a 4. Inoltre, dalla colonna *Blood pressure*, di tipo testuale, sono state estratte due nuove colonne, *Sistolica* e *Diastolica*, di tipo intero per separare le due tipologie di pressioni, migliorando il processo di analisi .

Figura 4.1: Creazione della nuova colonna *New Sleep Duration*

- **Eliminazione di colonne:** sono state eliminate tutte quelle colonne che non era possibile utilizzare ai fini dell’analisi dei dati come: *Sleep Duration* e *Blood Pressure*.

Per quanto riguarda la correzione o l’eliminazione di valori nelle celle non si è resa necessaria poichè il dataset non presentava errori rilevanti.

Una volta terminata la fase di trasformazione è iniziata quella di caricamento, dove i dati vengono trasferiti dall’editor di Power Query a Power BI Desktop.

4.2 Fase di analisi

Il seguente report Power BI si propone di analizzare il dataset ”*Sleep Health and Lifestyle*” relativo allo stile di vita e ai benefici del sonno. Dall’analisi sono emerse 4 dashboard, ognuna delle quali evidenzia il confronto tra uno specifico fattore e l’impatto che ha questo sul benessere psico-fisico delle persone. I fattori utilizzati per l’analisi sono i seguenti:

- **Genere:** nella prima dashboard vengono confrontati i dati sul benessere e sullo stile di vita fra i due generi, ponendo particolare attenzione al sonno e i suoi disturbi. L’obiettivo è comprendere come la durata e la qualità del sonno varino tra gli uomini e le donne e come possono influenzare il livello di stress percepito. Attraverso l’utilizzo di elementi visivi, vengono esplorate eventuali correlazioni tra i disturbi del sonno e gli altri parametri, al fine di individuare pattern significativi sulla gestione dello stress e sulla salute tra i due generi.
- **Età:** la seconda dashboard esplora l’impatto dell’età sulla salute dei soggetti, osservando come i parametri fisiologici evolvano nel tempo in relazione al sonno e allo stress. Grazie all’analisi di variabili come pressione sanguigna, frequenza cardiaca e qualità del sonno è possibile capire come l’invecchiamento influenzi la salute di un soggetto e, in particolare, se migliora o peggiora la qualità del riposo o se esista una relazione tra età e livello di stress.

- **Attività fisica:** la terza dashboard analizza gli effetti che l'attività fisica produce sul corpo umano, ponendo particolare attenzione ai risultati sullo stress percepito. Viene esplorata la relazione tra riposo e attività motoria, al fine di comprendere se sia in grado di prevenire eventuali disturbi del sonno.
- **Lavoro:** l'ultima dashboard si propone di analizzare l'impatto dell'attività lavorativo sullo stato di salute mentale e fisica di un soggetto. Attraverso l'analisi di parametri fisiologici si cerca di comprendere come le professioni sedentarie possano influire sulla salute fisica di un soggetto e se determinate categorie lavorative tendono ad essere più stressanti di altre. Inoltre, viene esaminato il legame con la qualità del sonno, al fine di comprendere se un'attività lavorativa sia maggiormente predisposta a disturbi del sonno.

Nelle prossime sottosezioni analizzeremo nel dettaglio gli strumenti di visualizzazione utilizzati al fine di evidenziare pattern significativi e relazioni tra variabili.

4.2.1 Analisi comparativa per genere



Figura 4.2: Analisi comparativa per genere

La prima pagina del report, in *Figura 4.2*, presenta cinque elementi visivi che mettono a confronto diversi attributi suddivisi per genere, evidenziando le principali differenze tra uomini e donne.

In alto a sinistra sono presenti due riquadri introduttivi, uno per ogni genere, che riportano il numero di persone e i valori medi delle ore di sonno giornaliere e della qualità del sonno. Dai dati emerge che, mediamente, le donne hanno una qualità del sonno maggiore rispetto agli uomini.

Alla loro destra è presente un istogramma a colonne raggruppate che mostra la distribuzione delle persone con disturbi del sonno, suddivise per genere. Dal grafico si può notare che, le donne sono più affette dall'apnea del sonno, nonostante presentino, in media, una qualità del sonno maggiore.

In basso a sinistra troviamo un riquadro contenente, in alto, un filtro a scorrimento che permette di classificare i dati per fasce di età. Sotto a questo, due grafici a linee mostrano la relazione tra il livello di stress, sull'asse x, e la media della qualità del sonno, nel primo, e la media delle ore di sonno nel secondo. Entrambi i grafici, categorizzati per genere, evidenziano un'andamento decrescente, da cui si può concludere una correlazione negativa tra stress e riposo.

Infine, in basso a destra, è presente un grafico a violino che rappresenta la distribuzione del livello di stress per uomini e donne. La forma dei violini risulta significativamente diversa tra i due generi, in particolare la media e la mediana del livello di stress degli uomini sono superiori a quelle delle donne. Questo ci suggerisce una diversa gestione dello stress fra i due generi.

4.2.2 Analisi dell'impatto dell'età su sonno, stress e parametri fisiologici

La seconda pagina del report, in *Figura 4.3*, analizza l'influenza dell'età sui parametri fisiologici, quali frequenza cardiaca e pressione sanguigna, e il loro rapporto con la qualità del sonno e il livello di stress. L'obiettivo è individuare eventuali pattern tra l'invecchiamento e il benessere psicofisico.

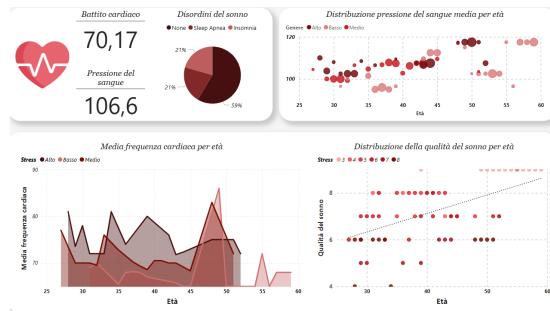


Figura 4.3: Analisi dell'impatto dell'età su sonno, stress e parametri fisiologici

In alto a sinistra sono presenti degli elementi visivi introduttivi che mostrano i valori medi del battito cardiaco e della frequenza cardiaca delle persone analizzate. Tali valori sono affiancati da un grafico a torta che mostra la distribuzione dei disturbi del sonno, con la possibilità di filtrare i dati per osservare l'incidenza dei disturbi sugli altri parametri presenti nel report. In alto a destra il grafico a dispersione espone la relazione tra l'età, (sull'asse delle ascisse) e la media tra la pressione sistolica e diastolica. Questa misura è stata calcolata utilizzando una formula DAX che somma i valori delle colonne *sistolica* e *diastolica*, dividendoli per due. I marcatori nel grafico sono colorati in base alla categoria di stress, con colori più scuri che indicano livelli di stress più elevati. Sebbene la pressione sanguigna tenda ad aumentare lievemente con l'età, dall'analisi del grafico non emergono correlazioni significative tra pressione sanguigna e stress.

In basso a sinistra è presente un grafico ad aree che mostra la relazione tra l'età e la frequenza cardiaca media, con le aree colorate in base alla categoria di stress. Dall'analisi emerge che, fino ai 45 anni, i soggetti appartenenti alla categoria di stress più elevata tendono a presentare una frequenza cardiaca mediamente più alta rispetto agli altri gruppi. Infine, nella parte inferiore destra della pagina, un ulteriore grafico a dispersione rappresenta la distribuzione della qualità del sonno (asse y) rispetto all'età (asse x), classificata per livello di stress. Dal grafico si può dedurre che, ad una bassa qualità del sonno corrispondono alte concentrazioni di stress. Inoltre, l'andamento crescente del grafico indica che con l'avanzare dell'età la qualità del sonno tende a migliorare, come lo si può notare per la maggior parte dei soggetti dai 50 anni in poi.

4.2.3 Analisi dell'impatto dell'attività fisica su sonno e stress

La terza pagina del report, in *Figura 4.4*, esamina la relazione tra l'attività motoria e il benessere psicofisico. L'obiettivo è individuare eventuali pattern attraverso l'analisi di parametri come i passi giornalieri, il tempo dedicato all'esercizio fisico e l'indice di massa corporea (BMI).

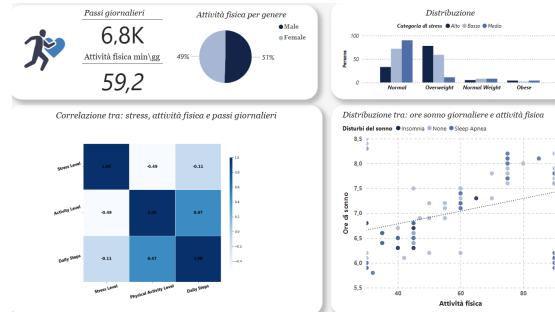


Figura 4.4: Analisi dell'impatto dell'attività fisica su sonno e stress

Nella parte superiore sinistra della pagina, sono presenti schede informative che mostrano i valori medi relativi all'attività fisica e al numero di passi compiuti quotidianamente dai soggetti analizzati. Queste grandezze sono accompagnate da un grafico a torta che illustra la distribuzione dell'attività fisica tra uomini e donne, permettendo anche di filtrare i parametri del report per genere.

Nella parte inferiore della pagina, è presente una *heatmap* che evidenzia la correlazione tra l'attività fisica, i passi giornalieri e il livello di stress. La mappa di calore è stata generata utilizzando la libreria **Seaborn** di Python. I coefficienti di correlazione calcolati non esprimono alcun rapporto tra i parametri, in quanto i valori sono vicini allo zero. Questo suggerisce che la relazione tra stress, attività fisica e passi giornalieri potrebbe essere influenzata da altri fattori.

Sul lato destro della pagina, un grafico a colonne raggruppate mostra la distribuzione delle categorie di *BMI* tra le persone, classificati per livello di stress. Come si può notare, i soggetti appartenenti alla categoria *normal* tendono ad averere livelli di stress inferiori rispetto a quelli della classe *overweight*. Questo suggerisce che, l'eccesso di grasso corporeo potrebbe essere correlato a livelli di stress più elevati.

L'ultimo elemento visivo, situato nella parte inferiore destra della pagina, è un grafico a dispersione che analizza la relazione tra le ore di sonno giornaliere, sull'asse delle ordinate, e l'attività fisica sull'asse delle ascisse. I marcatori sono colorati in base alla presenza di un disturbo del sonno, al fine di far emergere eventuali correlazioni tra l'esercizio fisico e il riposo. L'analisi evidenzia una tendenza crescente: la maggior parte dei soggetti che pratica quotidianamente almeno 60 minuti di attività motoria, tende a dormire circa 7 ore. Questo ci porta a presupporre che un adeguato livello di attività motoria favorisca il riposo.

4.2.4 Analisi del benessere psicofisico in un contesto lavorativo

L'ultima pagina del report, in *Figura 4.5*, è dedicata all'analisi dell'effetto che il lavoro sul benessere psicofisico, ponendo particolare attenzione alla relazione tra livelli di stress, sonno, frequenza cardiaca e passi giornalieri.

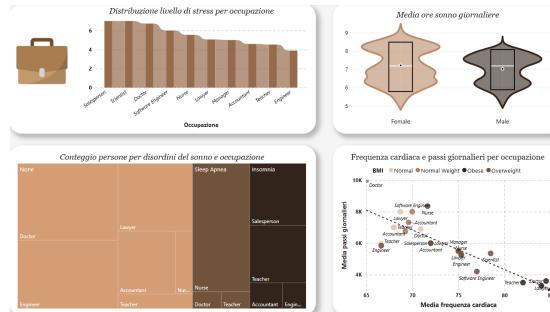


Figura 4.5: Analisi del benessere psicofisico in un contesto lavorativo

IL primo oggetto visivo posizionato in alto a sinistra presenta un grafico a nastri che esprime la distribuzione del livello di stress per categoria di lavoro. Come si può notare dal grafico, il lavoro sottoposto a maggiori livelli di stress è l'addetto alle vendite. Questo potrebbe essere dovuto da diversi fattori come la pressione nel raggiungimento degli obiettivi di vendita e orari di lavoro prolungati. Al contrario, gli ingegneri registrano livelli di stress mediamente più bassi.

Sul lato destro della pagina, un grafico a violino illustra la distribuzione del valore medio delle ore di sonno giornaliere suddivisa per genere. L'analisi mostra che la media e la mediana delle ore di sonno giornaliere sono simili per

entrambe le categorie. Anche la forma del violino, che rappresenta la densità della distribuzione, risulta piuttosto simile, questo suggerisce che le ore di sonno non sono particolarmente influenzate dal genere.

Nella parte inferiore a sinistra è presente una *mappa ad albero* che espone la distribuzione dei disordini del sonno per categorie di lavoro. Dall'analisi emerge che l'apnea del sonno è molto diffusa tra gli *infermieri*, questo potrebbe essere dovuto a ritmi di lavoro irregolari e turni di notte. Per quanto riguarda l'insonnia, risulta maggiormente diffusa tra gli addetti alle vendite e i professori; questo potrebbe essere causato da carichi di lavoro elevati che possono influire negativamente sul riposo notturno.

L'ultimo oggetto visivo del report è un grafico a dispersione che rappresenta la relazione tra la media della frequenza cardiaca e la media dei passi giornalieri. I punti sono colorati in base alla categoria di BMI e suddivisi per occupazione. Come si può osservare dal grafico i soggetti appartenenti alla categoria *obese* registrano una frequenza cardiaca elevata e un numero di passi giornalieri più basso, ciò potrebbe indicare una maggiore predisposizione a problemi cardio-vascolari, soprattutto se associati a lavori sedentari. Al contrario, le categorie *normal* e *normal weight* si posizionano maggiormente nel range tra i 6.000 e gli 8.000 passi giornalieri, con una frequenza cardiaca tra i 65 e i 75 battiti al minuto. Poiché in questa fascia sono presenti una varietà di lavori sedentari e non, mantenere un livello adeguato di attività fisica potrebbe contribuire a valori più equilibrati della frequenza cardiaca.