

UNIVERSITÀ DEGLI STUDI DI SALERNO



DIPARTIMENTO DI INFORMATICA

Corso di Laurea Magistrale in Informatica
Statistica e Analisi dei Dati

Green Growth Indicators

Valutazione delle Interazioni tra Crescita Economica e
Impatto Ambientale

Candidata:
Alessia Ture

Professori:
Prof. Stefano Cirillo
Prof. Luigi Di Biasi

ANNO ACCADEMICO 2023/2024

Indice

1	Introduzione	1
1.1	Contestualizzazione del Problema	1
1.2	Obiettivi dell'Analisi	3
2	Dataset	4
2.1	Origine e Caratteristiche del Dataset	4
2.1.1	Dimensione e Struttura	5
2.1.2	Periodo di Tempo	6
2.2	Pulizia e Pre-elaborazione dei Dati	6
2.2.1	Colonne selezionate	6
2.2.2	Variabili Selezionate	6
2.2.3	Paesi selezionati	8
2.3	Trattamento dei Dati Mancanti	9
2.3.1	Integrazione dei Dataset Aggiuntivi	10
2.3.2	Interpolazione Lineare	10
2.3.3	Imputazione Mediana	10
2.3.4	Specifiche delle Variabili con Dati Mancanti	10
3	Analisi Esplorativa	12
3.1	Obiettivi dell'analisi esplorativa	12
3.2	Analisi Univariata	12
3.2.1	Emissioni di CO ₂ basate sulla produzione	13
3.2.2	Emissioni di CO ₂ basate sulla domanda	19
3.2.3	Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale	26
3.2.4	Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili	32
3.2.5	Fornitura di energia rinnovabile, % del totale fornitura energetica	39
3.2.6	Elettricità Rinnovabile, % della Totale Produzione Elettrica	47
3.2.7	Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)	56
3.2.8	Percentuale di popolazione esposta a più di 10 µg/m ³	64
3.2.9	Costi del Benessere per le Mortalità Precoce dovute all'Esposizione a Particulate Matter 2.5 (PM2.5) Ambientali, Equivalente Prodotto Interno Lordo (PIL)	72
3.2.10	Area Marina Protetta % superficie terrestre	81
3.2.11	Area Terrestre Protetta % superficie terrestre	89
3.2.12	Tasse Correlate all' Ambiente, % del Prodotto Interno Lordo (PIL)	97

3.2.13	Prodotto Interno Lordo (PIL)	104
3.3	Diagramma di Pareto	110
3.3.1	Investimenti in Energie Rinnovabili	110
3.3.2	Gestione delle Risorse Idriche	112
3.3.3	Emissioni di CO ₂	113
3.4	Analisi Bivariata	115
3.4.1	Matrice di Correlazione	115
3.5	Scatterplots	116
3.5.1	Emissione di CO ₂ basate sulla produzione vs. Emissioni di CO ₂ basate sulle domande	117
3.5.2	Popolazione con accesso a acqua potabile migliorata e Prodotto Interno Lordo (PIL) reale pro capite	118
3.5.3	Mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5) e i Costi per il benessere delle mortalità premature	119
3.5.4	Energia Rinnovabile e Stress Idrico	120
3.5.5	Energia elettrica rinnovabile, % della generazione totale di elettricità e For- nitura di energia rinnovabile, % del totale dell'approvvigionamento energetico	121
3.5.6	Percentuale della popolazione esposta a più di 10 µg/m ³ e la Mortalità dovuta all'esposizione a PM2.5 ambientali	122
3.5.7	Costi per il benessere delle mortalità premature dovute a esposizione e la Percentuale della popolazione esposta a più di 10 µg/m ³	123
3.5.8	Emissioni di CO ₂ Basate sulla Domanda vs. Tasse Correlate all'Ambiente (% del Prodotto Interno Lordo (PIL))	124
3.5.9	Prodotto Interno Lordo (PIL) Reale Pro Capite vs. Mortalità da Esposizione a Particulate Matter 2.5 (PM2.5) Ambientali	125
3.5.10	Popolazione con Accesso a Acqua Potabile Migliorata (% della Popolazione Totale) vs. Percentuale di Popolazione Esposta a Più di 10 µg/m ³ :	126
3.5.11	Energia Elettrica Rinnovabile (Percentuale della Generazione Totale di Elet- tricità) vs. Percentuale di Popolazione Esposta a Più di 10 µg/m ³	127
4	Regressione	128
4.1	Emissioni Prodotte vs. Domandate	129
4.1.1	Risultati del Modello	129
4.1.2	Analisi dei Residui	130
4.2	Accesso all'Acqua Potabile e PM2.5	132
4.2.1	Risultati del Modello	132
4.2.2	Analisi dei Residui	133
4.3	Elettricità Rinnovabile e PM2.5	135
4.3.1	Risultati del Modello	135
4.3.2	Analisi dei Residui	136
4.4	Energia Rinnovabile e Stress Idrico	138
4.4.1	Risultati del Modello	138
4.4.2	Analisi dei Residui	139
4.5	Fornitura vs. Elettricità Rinnovabile	141
4.5.1	Risultati e Interpretazione	141
4.5.2	Analisi dei residui	142
4.6	Esposizione a PM2.5 e Mortalità	144

4.6.1	Modello di Regressione Multipla	144
4.6.2	Risultati del Modello	144
4.6.3	Analisi dei Residui	145
4.7	Conclusioni e Implicazioni Politiche	147
5	Analisi dei Cluster	148
5.1	Introduzione	148
5.2	Obiettivi dell'Analisi dei Cluster	149
5.3	Selezione delle variabili	149
5.4	Metodi Gerarchici	149
5.4.1	Metodo del legame singolo	149
5.4.2	Metodo del legame completo	151
5.4.3	Metodo del legame medio	152
5.4.4	Metodo della mediana	153
5.4.5	Metodo del centroide	155
5.4.6	Confronto dei Silhouette Score	156
5.4.7	Conclusione sui Metodi Gerarchici di Clustering	157
5.5	Metodi Non Gerarchici	158
5.5.1	Risultati K-Means con 2 Cluster	158
5.5.2	Risultati K-Means con 3 Cluster	159
5.5.3	Risultati K-Means con 4 Cluster	160
5.5.4	Risultati ottenuti	161
6	Inferenza Statistica	162
6.1	Stima Puntuale	162
6.1.1	Criterio del Chi-Quadrato	162
6.1.2	Metodo dei Momenti	166
6.2	Intervalli di Confidenza	167
6.2.1	Metodo Pivotal	167
6.3	Confronto tra Due Popolazioni	170
6.3.1	Verifica della Normalità	170
6.3.2	Intervallo di Confidenza per la Differenza tra le Medie	171
6.3.3	Risultati del Confronto tra le Medie delle Popolazioni	172
6.4	Verifica delle Ipotesi	173
6.4.1	Caso 1: Test Unilaterale Sinistro per le Emissioni di CO ₂	173
6.4.2	Caso 2: Test Unilaterale Destro per l'Energia Rinnovabile	174
7	Conclusioni	175

Capitolo 1

Introduzione

Questo documento presenta un'analisi approfondita dei dati relativi alle interazioni tra politiche ambientali, sostenibilità ambientale e sviluppo economico, utilizzando il dataset Organization for Economic Co-operation and Development (OECD). L'analisi mira a identificare tendenze, correlazioni e insight riguardanti l'impatto delle politiche ambientali sulle società e le economie a livello globale.

1.1 Contestualizzazione del Problema

L'importanza di un'analisi approfondita dei dati ambientali è divenuta cruciale nell'attuale contesto globale. Le sfide emergenti, come l'aumento dell'inquinamento, la gestione delle risorse idriche e gli impatti dei cambiamenti climatici, stanno influenzando significativamente il pianeta e la vita umana. In questo contesto, emerge la necessità di esaminare come l'equilibrio tra sviluppo economico, sostenibilità ambientale e benessere umano si stia evolvendo.

Le condizioni ambientali, influenzate da variabili complesse come l'inquinamento atmosferico, l'uso delle risorse idriche e le Politiche Ambientali, interagiscono in maniera significativa con le condizioni socio-economiche globali. Le decisioni politiche, le strategie aziendali e i comportamenti individuali sono fattori chiave nella definizione della salute dell'ambiente, che a sua volta impatta sulla qualità della vita.

La sostenibilità ambientale può essere definita come la capacità di un sistema di soddisfare le proprie esigenze attuali senza compromettere la capacità delle generazioni future di soddisfare le proprie. In altre parole, la sostenibilità ambientale implica la ricerca di un equilibrio tra lo sviluppo economico e la protezione dell'ambiente.

Le Politiche Ambientali possono essere classificate in base a diversi criteri. Una possibile classificazione si basa sul tipo di obiettivo che le politiche perseguono. In questo caso possono essere classificate in:

- Le *politiche di prevenzione* che mirano a ridurre le cause dell'inquinamento e del degrado ambientale.
- Le *politiche di mitigazione* che mirano a ridurre gli impatti negativi delle attività umane sull'ambiente.

- Le *politiche di adattamento* che mirano a ridurre la vulnerabilità dei sistemi sociali e naturali agli impatti dei cambiamenti climatici.

Un'altra possibile classificazione delle Politiche Ambientali si basa sul settore di intervento. In questo caso, le politiche ambientali possono essere classificate in politiche ambientali per l'energia, per l'industria, per l'agricoltura, per i trasporti, ecc.

Evidenze scientifiche suggeriscono che l'adozione di Politiche Ambientali può portare a una serie di benefici, tra cui:

1. Miglioramento della qualità dell'aria e dell'acqua
2. Riduzione dell'inquinamento e del degrado ambientale
3. Protezione della Biodiversità
4. Creazione di opportunità economiche

Tuttavia, l'adozione di Politiche Ambientali può anche comportare alcuni costi, come:

1. Aumento dei costi di produzione
2. Perdita di posti di lavoro
3. Aumento dei prezzi dei beni e dei servizi

L'inquinamento atmosferico e idrico può avere un impatto negativo sulla salute umana e sull'ambiente. L'adozione di Politiche Ambientali, come le normative sulle emissioni e la gestione delle acque reflue, può aiutare a migliorare la qualità dell'aria e dell'acqua, riducendo i rischi per la salute e aumentando la qualità della vita. L'inquinamento e il degrado ambientale possono causare danni significativi all'ambiente, come la perdita di habitat, l'erosione del suolo e l'acidificazione degli oceani. L'adozione di Politiche Ambientali, come le leggi sulla protezione della fauna selvatica e la conservazione del suolo, può aiutare a ridurre l'inquinamento e il degrado ambientale, proteggendo l'ambiente e le risorse naturali.

La Biodiversità è essenziale per la salute dell'ecosistema e per il benessere umano. L'adozione di Politiche Ambientali, come le leggi sulla conservazione della natura e l'uso sostenibile delle risorse, può aiutare a proteggere la biodiversità, garantendo la sopravvivenza delle specie e dei loro habitat.

Le Politiche Ambientali possono creare opportunità economiche, creando nuovi posti di lavoro e stimolando l'innovazione. Ad esempio, le politiche che promuovono l'energia rinnovabile e l'efficienza energetica possono creare posti di lavoro nei rispettivi settori.

L'attuazione di Politiche Ambientali può influire sui prezzi dei beni e servizi. Ad esempio, normative sulle emissioni possono aumentare i costi di produzione dell'energia, con conseguente possibile aumento dei prezzi dell'energia per i consumatori.

Questo studio si prefigge di esaminare in modo dettagliato queste dinamiche complesse, con un focus particolare sulle Politiche Ambientali e il loro effetto diretto sull'ambiente, sulla popolazione e sull'economia. L'obiettivo è di comprendere in che modo l'adozione di politiche sostenibili influisca sulla gestione delle risorse naturali, sulla qualità dell'aria e dell'acqua, e sulle opportunità economiche.

1.2 Obiettivi dell'Analisi

Il presente progetto si propone di raggiungere una serie di obiettivi chiari e ben definiti al fine di esplorare e comprendere le dinamiche legate alle variabili ambientali, economiche e di salute considerate nel dataset. Gli obiettivi dell'analisi sono i seguenti:

1. **Analisi Temporale delle Emissioni di CO₂:** Esaminare l'andamento nel tempo delle emissioni di CO₂ per identificare trend, fluttuazioni e possibili cambiamenti nel corso degli anni.
2. **Valutazione dell'Impatto delle Politiche Ambientali:** Determinare se l'adozione di Politiche Ambientali è associata a variazioni significative nelle emissioni di CO₂ e nelle altre variabili ambientali considerate nel dataset.
3. **Correlazioni tra Emissioni di CO₂ ed Economia:** Analizzare l'effetto delle emissioni di CO₂ sull'economia nazionale.
4. **Effetto delle Energie Rinnovabili sulle Emissioni:** Verificare se l'incremento dell'uso delle energie rinnovabili è associato a una significativa riduzione delle emissioni di CO₂.
5. **Impatto dell'Accesso all'Acqua Potabile sulla Salute:** Esaminare se l'accesso a fonti di acqua potabile ha un effetto significativo sulla salute della popolazione e valutare come l'inquinamento impatta l'accesso alle acque.
6. **Esposizione a PM2.5 e Costi Correlati:** Valutare l'effetto dell'esposizione a PM2.5 sulla salute pubblica e quantificare i costi associati alle morti premature dovute all'inquinamento atmosferico.
7. **Confronto tra Paesi con Diverse Emissioni di CO₂:** Confrontare le emissioni di CO₂ tra paesi con diversi livelli di reddito pro capite al fine di individuare eventuali differenze significative.
8. **Analisi di Impatto Ambientale:** Valutare l'impatto delle emissioni di CO₂ e dell'uso di risorse naturali sullo Stress Idrico e sulle aree protette.

L'obiettivo finale è valutare l'impatto complessivo delle Politiche Ambientali sulla sostenibilità economica ed ecologica, esplorando le interazioni tra le variabili considerate.

Capitolo 2

Dataset

2.1 Origine e Caratteristiche del Dataset

Il dataset scelto per questa analisi proviene dall'Organization for Economic Co-operation and Development (OECD) e viene denominato “*OECD Green Growth*” che comprende una serie di indicatori progettati per monitorare i progressi verso una crescita verde, supportando così la formulazione di politiche e informando il pubblico.

Il “Organization for Economic Co-operation and Development (OECD) Green Growth” aggrega dati e indicatori in una vasta gamma di domini, attingendo da diverse basi di dati dell'OECD e da fonti esterne. Include paesi membri e candidati dell'OECD, partner chiave e altri paesi non OECD selezionati.

Gli indicatori sono raggruppati in quattro categorie principali:

1. **Produttività ambientale e delle risorse:** indica se la crescita economica sta diventando più verde con un uso più efficiente del capitale naturale.
2. **Base di attività naturali:** indica i rischi per la crescita derivanti da una base di attività naturali in declino.
3. **Dimensione ambientale della qualità della vita:** indica come le condizioni ambientali influenzano la qualità della vita e il benessere delle persone.
4. **Opportunità economiche e risposte politiche:** indica l'efficacia delle politiche nel promuovere la crescita verde e descrive le risposte della società necessarie per garantire opportunità di lavoro e di impresa.

2.1.1 Dimensione e Struttura

Il dataset consiste in 411887 righe e 15 colonne. Ogni riga rappresenta un anno specifico per un determinato paese.

Le colonne del dataset sono le seguenti:

1. **COU**: rappresenta il codice identificativo del paese. Ogni paese ha un codice unico che lo distingue dagli altri.
2. **Country**: rappresenta il nome completo del paese. Ad esempio, “AUS” nella colonna “COU” potrebbe corrispondere a “Australia” nella colonna “Country”.
3. **VAR**: un codice identificativo per la variabile specifica all’interno del dataset.
4. **Variable**: descrizione dettagliata della variabile rappresentata nella colonna “VAR”.
5. **YEA**: indica il codice univoco dell’anno a cui si riferiscono i dati nella riga corrispondente.
6. **Year**: indica l’anno a cui si riferiscono i dati nella riga corrispondente.
7. **Unit Code**: Un codice che rappresenta l’unità di misura utilizzata per la variabile.
8. **Unit**: una descrizione completa dell’unità di misura rappresentata nella colonna “Unit Code”.
9. **PowerCode Code**: un codice per una potenza o un moltiplicatore utilizzato per esprimere i valori della variabile. Ad esempio, un valore potrebbe essere espresso in “migliaia” o “milioni”.
10. **PowerCode**: Una descrizione dettagliata del codice di potenza o moltiplicatore rappresentato nella colonna “PowerCode Code”.
11. **Reference Period Code**: Un codice che indica un periodo di riferimento specifico per i dati. Ad esempio, i dati potrebbero riferirsi a un anno fiscale o a un anno civile.
12. **Reference Period**: Una descrizione dettagliata del periodo di riferimento rappresentato nella colonna “Reference Period Code”.
13. **Value**: Il valore numerico o dato per la variabile specificata nella riga corrispondente. Ad esempio, potrebbe rappresentare le emissioni di CO₂ di un paese in un determinato anno.
14. **Flag Codes**: Un codice che fornisce informazioni aggiuntive sul dato. Ad esempio, un codice potrebbe indicare se un dato è stimato o provvisorio.
15. **Flags**: Una descrizione dettagliata del codice flag rappresentato nella colonna “Flag Codes”.

I possibili valori per il campo flags sono:

- *B: Break in series*
- *C: Confidential*
- *E: Estimated*
- *I: Incomplete*
- *n: National estimate*

2.1.2 Periodo di Tempo

Il dataset copre un arco temporale che va dall'anno 1990 all'anno 2022. Questo ampio intervallo temporale consente di effettuare analisi dettagliate sulle tendenze a lungo termine in ambito ambientale ed economico.

2.2 Pulizia e Pre-elaborazione dei Dati

Prima di procedere con l'analisi, i dati sono stati sottoposti a una fase di pulizia e pre-elaborazione. Questo ha incluso la gestione dei dati mancanti, la correzione di eventuali anomalie e la normalizzazione delle variabili per garantire coerenza e affidabilità nelle analisi successive.

2.2.1 Colonne selezionate

Non tutte le colonne presenti nel dataset sono rilevanti per l'analisi proposta, pertanto è stata effettuata una selezione mirata delle colonne da includere. La scelta delle colonne è stata guidata dal principio di mantenere solo quelle informazioni essenziali per rispondere efficacemente agli obiettivi di ricerca. Le colonne selezionate sono:

- **Country:** Questa colonna è stata mantenuta perché fornisce una chiara identificazione dei paesi, essenziale per qualsiasi analisi geografica o comparativa.
- **Variable:** Fondamentale per comprendere la natura specifica dei dati in ogni riga, questa colonna ci dice esattamente cosa stiamo misurando.
- **YEA:** L'anno è una dimensione temporale fondamentale, permettendo analisi di serie storiche e trend nel tempo.
- **Unit:** Queste colonne sono state mantenute per garantire che ogni dato possa essere interpretato correttamente nel suo contesto. Sapere l'unità di misura è cruciale per comprendere e interpretare i dati correttamente.
- **Value:** Ovviamente, il valore effettivo dei dati è fondamentale per qualsiasi analisi.

2.2.2 Variabili Selezionate

Le variabili scelte coprono un'ampia gamma di aspetti legati all'ambiente, alla salute, all'economia e alle risorse naturali. Di seguito viene fornita una breve descrizione e giustificazione per ciascuna categoria di variabili selezionate:

1. Emissioni di CO₂

- **Production-based CO₂ emissions:** Misura le emissioni di CO₂ generate dalla produzione di beni e servizi in un paese. È fondamentale per valutare l'impatto ambientale diretto delle attività produttive.
- **Demand-based CO₂ emissions:** Rappresenta le emissioni di CO₂ associate alla domanda di beni e servizi, incluse le importazioni. Questo dato è utile per comprendere l'impronta carbonica complessiva legata al consumo.

2. Energia

- **Renewable energy supply, % total energy supply:** Indica la percentuale di energia proveniente da fonti rinnovabili sul totale dell'approvvigionamento energetico. È un indicatore chiave per valutare la transizione energetica verso fonti più sostenibili.
- **Renewable electricity, % total electricity generation:** Rappresenta la quota di elettricità generata da fonti rinnovabili rispetto al totale della produzione elettrica. È cruciale per monitorare i progressi nella decarbonizzazione del settore elettrico.

3. Acqua

- **Population with access to improved drinking water sources, % total population:** Misura la percentuale della popolazione che ha accesso a fonti di acqua potabile migliorate, un indicatore essenziale della qualità della vita e della salute pubblica.
- **Water stress, total freshwater abstraction as % total available renewable resources:** Indica il livello di Stress Idrico calcolato come percentuale di prelievo d'acqua rispetto alle risorse rinnovabili disponibili, importante per valutare la sostenibilità dell'uso delle risorse idriche.

4. Salute

- **Mortality from exposure to ambient PM2.5:** Misura la mortalità associata all'esposizione a particelle fini (PM2.5) nell'aria, un indicatore cruciale per la salute pubblica.
- **Percentage of population exposed to more than 10 micrograms/m³:** Rappresenta la percentuale della popolazione esposta a livelli di PM2.5 superiori ai limiti raccomandati, importante per valutare l'impatto dell'inquinamento atmosferico sulla popolazione.
- **Welfare costs of premature mortalities from exposure to ambient PM2.5, GDP equivalent:** Stima i costi economici delle morti premature dovute all'esposizione a PM2.5, espressi come percentuale del PIL, evidenziando l'impatto economico dell'inquinamento atmosferico.

5. Economia

- **Environmentally related taxes, % GDP:** Indica il livello di tassazione legata all'ambiente come percentuale del PIL, riflettendo l'impegno governativo nella promozione di politiche ambientali.
- **Real GDP per capita:** Misura il PIL reale pro capite, un indicatore chiave del benessere economico di una nazione.

6. Risorse Naturali

- **Terrestrial protected area, % land area:** Indica la percentuale del territorio nazionale che è classificato come area protetta, un indicatore della conservazione delle risorse terrestri.
- **Marine protected area, % total exclusive economic zone:** Misura la percentuale della zona economica esclusiva totale di un paese che è protetta, un indicatore dell'impegno nella conservazione degli ecosistemi marini.

2.2.3 Paesi selezionati

Il dataset originale includeva dati relativi a 236 paesi. Tuttavia, è stata effettuata una selezione mirata di 33 paesi per garantire che l'analisi fosse rappresentativa, inclusiva e rilevante rispetto alle diverse sfide globali legate alla crescita verde e alla sostenibilità ambientale. La selezione dei paesi è stata guidata dai seguenti criteri:

- **Assenza di valori nulli:** presenza di valori per le variabili scelte.
- **Rappresentatività Globale:** Inclusione di paesi da diverse regioni del mondo per garantire una prospettiva globale.
- **Diversi Livelli di Sviluppo Economico:** Selezione di paesi sviluppati, in via di sviluppo e paesi emergenti per riflettere una gamma di realtà economiche.
- **Disponibilità dei Dati:** Priorità a paesi con dati completi e affidabili nel database.
- **Diversità Ambientale e Sociale:** Considerazione di paesi con diverse sfide e politiche ambientali.

I paesi selezionati per l'analisi sono stati suddivisi per continente al fine di facilitare una comprensione più dettagliata delle analisi regionali. Di seguito è riportata la mappatura dei paesi secondo le rispettive subregioni:

1. *Europa Occidentale:* Germania, Francia, Regno Unito, Italia, Spagna, Paesi Bassi, Belgio, Lussemburgo.
2. *Europa Orientale:* Polonia, Ungheria, Danimarca, Finlandia, Grecia, Irlanda, Portogallo, Norvegia, Svezia.
3. *America:* Stati Uniti, Canada, Messico, Brasile, Argentina.
4. *Asia:* Giappone, Corea, Pakistan, Cina, India, Russia, Indonesia.
5. *Oceania e Africa:* Nuova Zelanda, Australia, Sud Africa, Nigeria.

Dopo questa prima fase il dataset contiene: 8317 righe e 5 colonne.

2.3 Trattamento dei Dati Mancanti

Inizialmente, sono stati identificati i valori nulli presenti nelle colonne del dataset, come mostrato nella Tabella 2.1.

Colonna	Valori Nulli
Paese	0
Variabile	0
Anno (YEA)	0
Valore	0
Unità	2,125

Tabella 2.1: Valori nulli per ciascuna colonna

I valori nulli sono presenti unicamente nella colonna `Unit` e sono stati sostituiti con il valore "Non specificato" per garantire la coerenza del dataset. Dopo questo trattamento, il dataset risultante contiene 8,317 righe e 5 colonne.

Il numero di righe atteso, considerando una serie temporale dal 1990 al 2021 per ciascuna delle 13 variabili selezionate in 33 paesi, sarebbe $31 \times 13 \times 32 = 12,896$. La discrepanza rispetto al numero attuale di righe suggerisce la presenza di dati mancanti o incompleti nel dataset.

Un primo segnale di dati mancanti è dato dal fatto che non tutte le variabili coprono l'intero periodo di osservazione. Ad esempio, la serie storica delle **Emissioni di CO₂ basate sulla produzione** ha 32 osservazioni per ciascun paese, mentre le **Emissioni di CO₂ basate sulla domanda** ne hanno solo 24, poiché le osservazioni arrivano fino al 2018. Per uniformare le analisi, si è deciso di portare tutte le serie storiche a un periodo dal 1995 al 2018.

L'Interpolazione Lineare è stata utilizzata per trattare i valori mancanti all'interno delle serie temporali. Tuttavia, l'interpolazione può essere applicata solo quando ci sono almeno due punti consecutivi non nulli. Pertanto, alcune variabili presentano ancora dati mancanti, come mostrato nella Tabella 2.2, per queste variabili, sono stati cercati dataset aggiuntivi per integrare i dati mancanti.

Variabile	Valori Mancanti
Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale	6
Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili	3
Tasse legate all'ambiente, % del PIL	2
Area marina protetta, % della zona economica esclusiva totale	2

Tabella 2.2: Riassunto dei valori mancanti per variabile

2.3.1 Integrazione dei Dataset Aggiuntivi

Per trattare il problema dei dati mancanti, sono stati integrati due dataset aggiuntivi per le variabili:

- **Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili.**
- **Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale.**

Questi dataset sono stati recuperati dal World Bank Data (<https://data.worldbank.org>). Poiché i dataset aggiuntivi contenevano solo una serie storica parziale, è stato necessario applicare un'ulteriore fase di interpolazione e imputazione.

2.3.2 Interpolazione Lineare

L'Interpolazione Lineare è stata utilizzata per gestire i dati mancanti nelle serie temporali. Questo metodo, semplice e intuitivo, è adatto quando i cambiamenti tra i punti dati consecutivi sono relativamente piccoli e lineari. Tuttavia, l'interpolazione lineare non può prevedere valori al di fuori dell'intervallo dei dati esistenti (nessuna estrapolazione), il che può limitare la sua applicabilità in alcuni casi.

2.3.3 Imputazione Mediana

Per quei paesi e variabili che, nonostante l'interpolazione, presentavano ancora dati mancanti, si è ricorso all'Imputazione Mediana. Questo metodo è stato utilizzato in particolare per i paesi che non avevano valori iniziali sufficienti per l'interpolazione.

2.3.4 Specifiche delle Variabili con Dati Mancanti

Di seguito sono riportate le specifiche per ciascuna variabile con dati mancanti e il metodo utilizzato per l'imputazione:

1. **Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale:** Anche dopo l'integrazione con il nuovo dataset, sono rimasti dati mancanti per i paesi *Cina, India, Sud Africa, Argentina e Australia*. Per ciascuno di questi paesi, sono stati utilizzati dati da paesi simili:
 - **Cina:** Dati interpolati dalla serie storica dell'Indonesia.
 - **India:** Dati interpolati dalla serie storica del Brasile.
 - **Sud Africa:** Dati interpolati dalla serie storica del Messico.
 - **Argentina:** Dati interpolati dalla serie storica del Messico e del Brasile.
 - **Australia:** Dati interpolati dalla serie storica degli Stati Uniti.
2. **Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili:** Questo dataset è stato integrato completamente, rimpiazzando 389 valori nulli.
3. **Tasse legate all'ambiente, % del PIL:** Per la Russia e l'Indonesia, sono stati utilizzati modelli di serie temporali (ARIMA) basati sulle serie storiche di paesi simili:

- **Russia:** Dati imputati usando la serie storica del Brasile.
 - **Indonesia:** Dati imputati usando la serie storica dell'India.
4. **Area Marina Protetta, % della zona economica esclusiva totale:** I paesi *Lussemburgo* e *Ungheria*, essendo senza sbocchi al mare, non hanno una zona economica esclusiva (EEZ) marina. Pertanto, si è ragionevolmente assunto che la percentuale di aree marine protette rispetto alla loro EEZ sia pari a zero.

Capitolo 3

Analisi Esplorativa

3.1 Obiettivi dell'analisi esplorativa

L'**Analisi Esplorativa dei Dati**, comunemente nota come EDA, rappresenta il pilastro fondamentale del processo di analisi dei dati. Questa fase iniziale ha l'obiettivo cruciale di consentire agli analisti e ai ricercatori di familiarizzare con i dati, comprendendone le caratteristiche principali, le tendenze, i pattern e, se presenti, le anomalie.

Attraverso l'EDA, esploriamo e visualizziamo i dati in modo non strutturato per scoprire i primi **insight** e le strutture intrinseche. Questo può includere l'identificazione di potenziali Correlazione tra variabili, la rilevazione di Outlier, la comprensione della Distribuzione dei Dati e la scoperta di eventuali errori o inconsistenze.

Inoltre, l'EDA è un processo iterativo e non lineare che spesso porta a ulteriori domande e ipotesi. Con una solida EDA, gli analisti possono formulare **ipotesi informate** e scegliere le tecniche statistiche o i modelli di ML più adatti per approfondire l'analisi.

La capacità di eseguire efficacemente un'EDA richiede non solo una conoscenza tecnica degli strumenti e delle tecniche analitiche ma anche un'acuta sensibilità per cogliere le **sfumature** e i **dettagli nascosti** nei dati. Un'EDA ben eseguita pone le basi per tutti i passaggi successivi dell'analisi dei dati, garantendo che le **decisioni** e le **previsioni** siano basate su una comprensione accurata e approfondita del dataset a disposizione.

3.2 Analisi Univariata

L'**Analisi Univariata** rappresenta il primo passo della Exploratory Data Analysis (EDA) ed è focalizzata sull'esame delle singole variabili nel dataset, senza considerare le relazioni con altre variabili. L'obiettivo principale dell'**Analisi Univariata** è comprendere le **caratteristiche**, la **distribuzione** e le **tendenze intrinseche** di ciascuna variabile.

3.2.1 Emissioni di CO₂ basate sulla produzione

Le emissioni di CO₂ basate sulla produzione sono espresse in milioni di tonnellate metriche. Sono incluse le emissioni di CO₂ dalla combustione di carbone, petrolio, gas naturale e altri combustibili. Le stime delle emissioni di CO₂ sono ottenute dal database IEA delle emissioni di CO₂ da combustione di combustibili. Le stime si basano sulle Emissioni territoriali.

Emissioni di CO ₂ basate sulla produzione	
Media	700
Mediana	303
Simmetria	3.821358
Curtosi	15.51168
Frequenza assoluta	744
Massimo	9777
Minimo	7.12
Deviazione standard	1451.6
Coefficiente di variazione	2.07494
Varianza	2107143
Q1	56.4
Q3	502

Tabella 3.1: Statistiche descrittive per la variabile: **Emissioni di CO₂ basate sulla produzione**

Il **Boxplot** nella Figura 3.1 offre una rappresentazione visiva della distribuzione delle emissioni di CO₂ basate sulla produzione durante il periodo preso in considerazione. Questo grafico fornisce importanti informazioni statistiche: La **mediana** delle emissioni è di 303 milioni di tonnellate metriche, indicando che il valore centrale delle emissioni è relativamente basso rispetto alla gamma totale.

La **media**, invece, è di 700 milioni di tonnellate metriche, che è significativamente più alta della mediana. Questo suggerisce la presenza di valori estremi superiori che influenzano la media e indicano una distribuzione con una coda lunga verso valori più alti. Il boxplot mostra chiaramente la presenza di alcuni **valori estremi superiori**, evidenziati come punti al di fuori dei "baffi" del grafico. Ad esempio, il valore massimo delle emissioni raggiunge 9777 milioni di tonnellate metriche, suggerendo la presenza di alcune nazioni con emissioni eccezionalmente elevate. Il **valore minimo** mostra che c'è almeno un paese con emissioni di CO₂ molto basse, che suggerisce una vasta gamma di profili di emissione tra i diversi paesi. Una **deviazione standard elevata** indica un'ampia variazione nelle emissioni di CO₂ tra i paesi. Questa significativa dispersione dei dati rispecchia l'eterogeneità nelle pratiche industriali e nelle politiche ambientali a livello globale.

Questo boxplot è utile per comprendere la variabilità delle emissioni di CO₂ tra i paesi e identificare i casi in cui sono necessarie azioni specifiche per affrontare le emissioni elevate e promuovere la sostenibilità ambientale.

Il **coefficiente di variazione** è una misura standardizzata della dispersione di una distribuzione. È il rapporto tra la deviazione standard e la media. Un coefficiente di variazione di 2.07 suggerisce un'alta variabilità relativa nella distribuzione delle emissioni di CO₂. Indica che le emissioni variano considerevolmente tra diversi paesi.

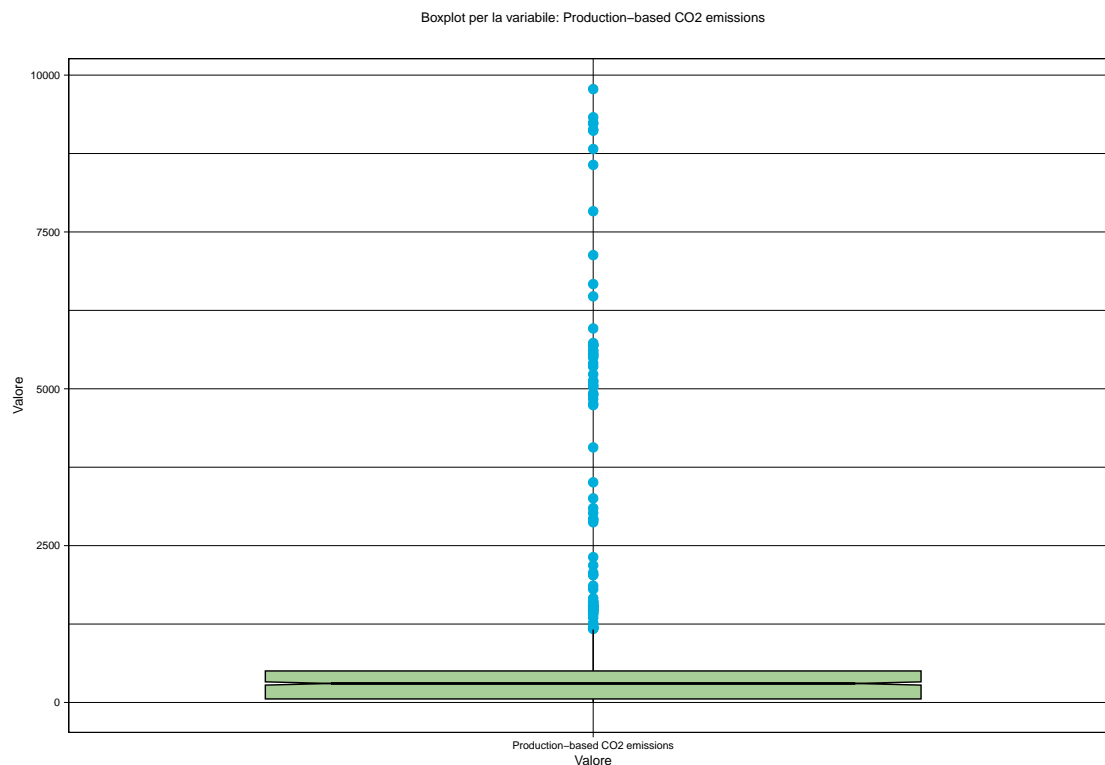


Figura 3.1: Boxplot delle emissioni di CO₂ basate sulla produzione.

Il **Primo Quartile (Q1)** e il **Terzo Quartile (Q3)** mostrano che il 50% delle emissioni dei paesi si trova in questo intervallo. Un Q1 basso e un Q3 relativamente alto indicano una distribuzione con molte osservazioni concentrate verso il valore più basso e poche con valori molto elevati.

La **simmetria positiva** della distribuzione, indicata da $\text{skewness} = 3.82$, conferma la presenza di una coda lunga verso valori più alti. Ciò significa che alcune nazioni hanno un impatto ambientale molto più significativo rispetto ad altre, influenzando la distribuzione complessiva delle emissioni. In termini pratici, indica la presenza di un numero relativamente piccolo di paesi con emissioni eccezionalmente elevate, che distorcono la media verso l'alto. Una **curtosi** di 15.51 è significativamente alta, suggerendo la presenza di valori anomali estremi nelle emissioni di CO₂. Ciò indica che ci sono paesi che emettono quantità di CO₂ molto maggiori rispetto alla norma, influenzando la distribuzione complessiva.

L'**istogramma** nella Figura 3.2 mostra la frequenza delle emissioni di CO₂ basate sulla produzione in milioni di tonnellate metriche. Ogni barra rappresenta l'intervallo di emissioni e la loro frequenza relativa tra i paesi. La forma dell'istogramma evidenzia la **distribuzione asimmetrica** delle emissioni, con una maggioranza di paesi che hanno emissioni relativamente basse e pochi paesi con valori estremamente elevati. Questo è coerente con l'elevata simmetria e curtosi osservate, indicando una distribuzione distorta con coda pesante verso valori più alti.

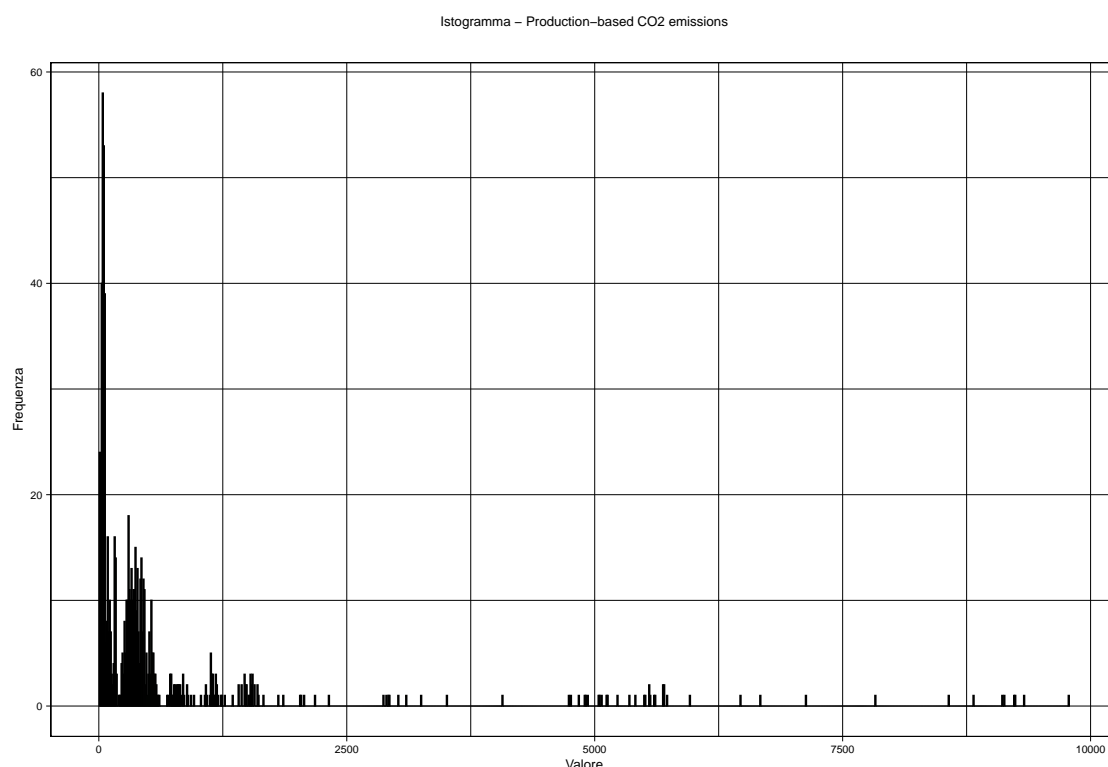


Figura 3.2: Istogramma delle emissioni di CO₂ basate sulla produzione.

La **funzione di densità** nella Figura 3.3 fornisce una rappresentazione continua della distribuzione probabilistica delle emissioni di CO₂. La curva indica come si distribuiscono le probabilità per differenti livelli di emissioni. La curva mostra chiaramente un picco a bassi livelli di emissioni, con una lunga coda verso destra. Questa coda rappresenta la presenza di valori estremi, coerente con una distribuzione asimmetrica e con una significativa presenza di outliers. Questo suggerisce che, mentre la maggior parte dei paesi ha emissioni basse o moderate, esiste un piccolo numero di paesi con emissioni molto elevate, influenzando la distribuzione complessiva.

Di seguito, nelle figure 3.5, 3.4, 3.6, 3.7, 3.8 sono riportate le serie storiche per la variabile. Sull'asse orizzontale poniamo gli anni dal 1995 al 2018, mentre sull'asse verticale sono indicate le tonnellate di CO₂ prodotte in (United States Dollar (USD)/kg) in **scala logaritmica** per rendere il grafico più leggibile, in quanto i paesi che si trovano negli stessi continenti tendono ad avere un andamento simile.

La Figura 3.4 mostra l'evoluzione temporale delle emissioni di CO₂ basate sulla produzione per **i paesi del continente Americano**. La scala logaritmica sull'asse verticale permette una visualizzazione più chiara delle variazioni su un ampio intervallo di valori. Si osserva una tendenza generale alla stabilità o a un lieve aumento delle emissioni in molti paesi, con l'Argentina che mostra una crescita notevole. Questo può essere interpretato come un riflesso dell'aumento dell'industrializzazione in alcuni paesi del continente.

La Figura 3.5 presenta le serie storiche delle emissioni di CO₂ per **i paesi asiatici**. La scala logaritmica evidenzia sia le tendenze generali sia le variazioni annuali. Si nota una rapida crescita

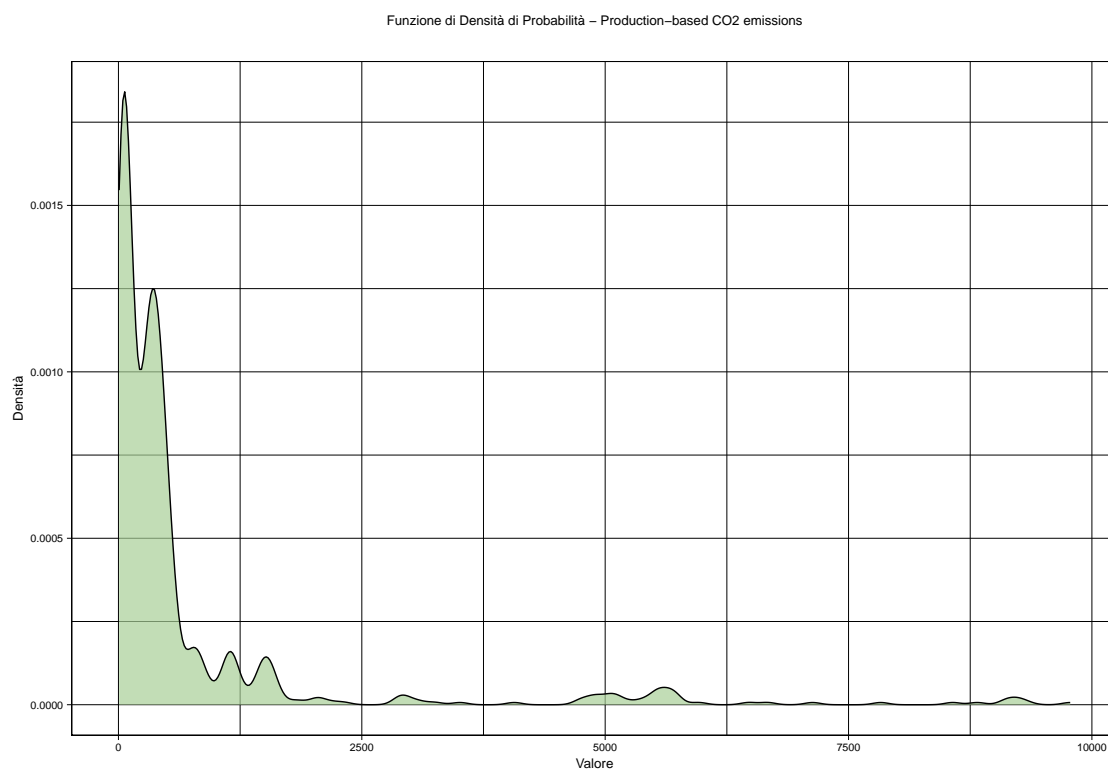
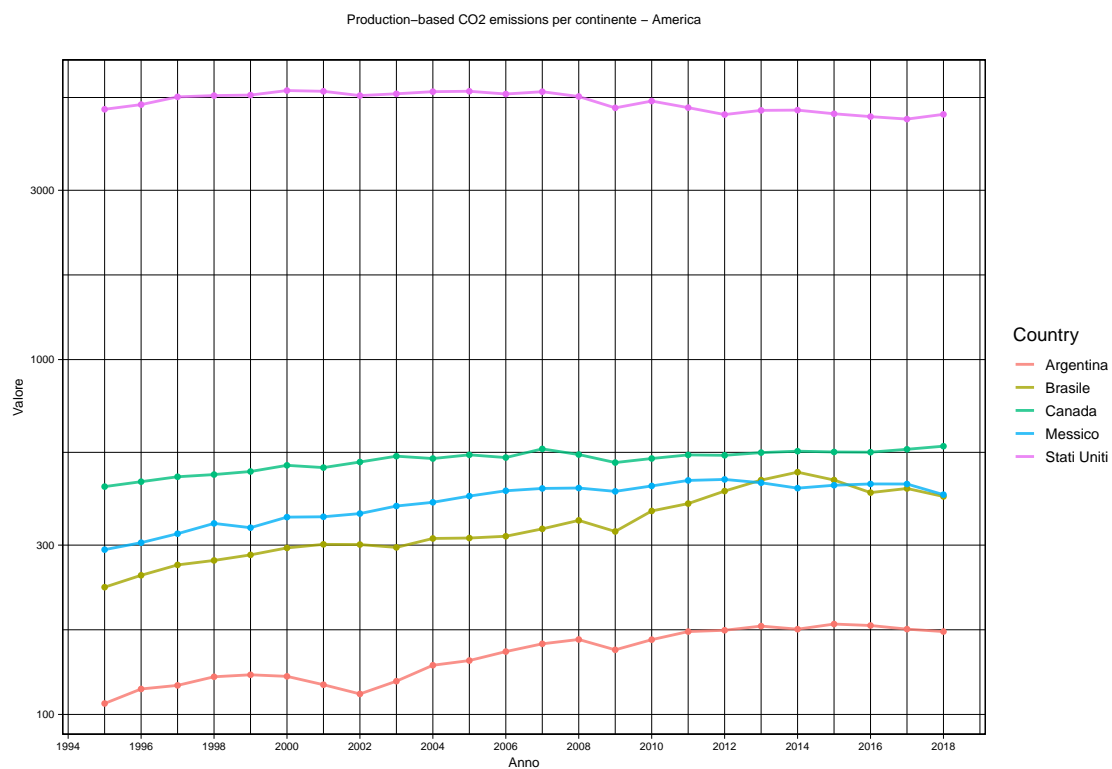
Figura 3.3: Funzione di densità delle emissioni di CO₂ basate sulla produzione.

Figura 3.4: Serie storica per il continente Americano.

delle emissioni in Cina, Indonesia e India, probabilmente dovuta alla rapida industrializzazione. Giappone e Russia mostrano una diminuzione temporanea delle emissioni intorno al 2008-2009, forse a causa della recessione economica, seguita da una successiva ripresa.

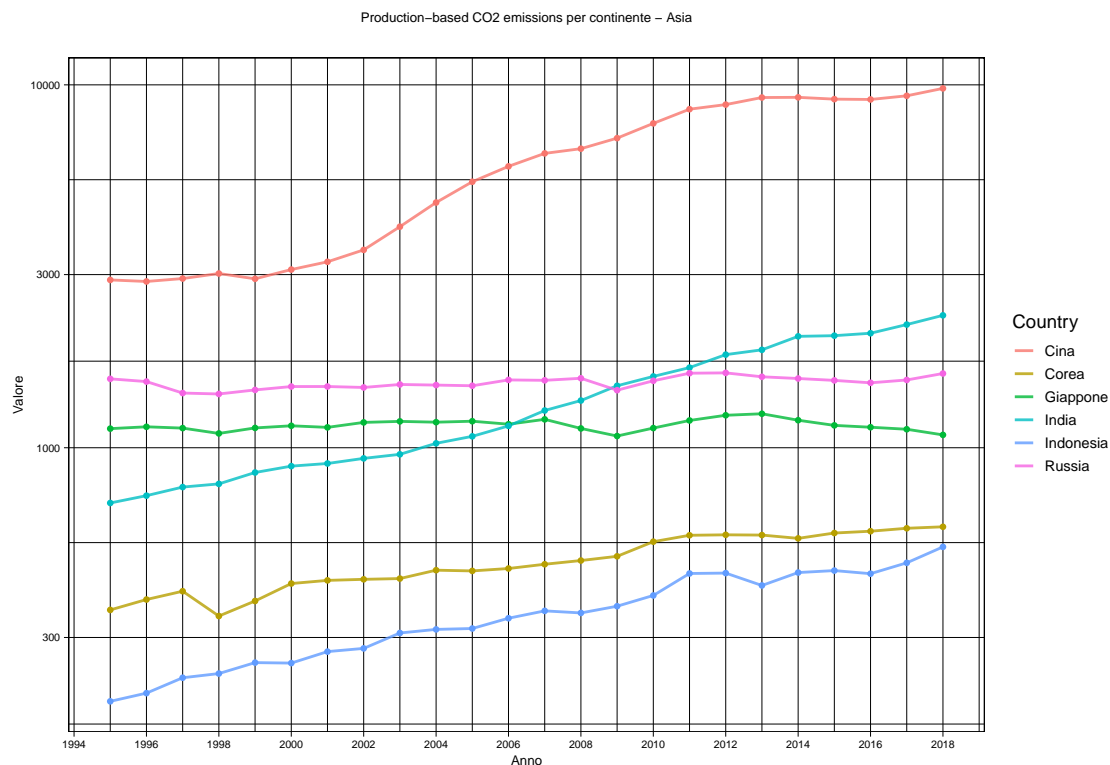


Figura 3.5: Serie storica per il continente Asiatico.

Il grafico 3.6 illustra le emissioni di CO₂ negli anni per i **paesi dell'Europa Occidentale**, usando una scala logaritmica per una migliore leggibilità. In generale, i paesi europei mostrano emissioni relativamente stabili. Paesi come la Grecia, la Danimarca, la Finlandia e la Svezia hanno visto un calo delle emissioni, riflettendo le forti politiche ambientali e, nel caso della Grecia, gli effetti della recessione economica.

Il grafico 3.7 rappresenta le emissioni di CO₂ nel tempo per **l'Europa Orientale**. La scala logaritmica aiuta a delineare le tendenze in un contesto di variazione significativa. Le tendenze mostrano una variabilità maggiore rispetto ai paesi dell'Europa Occidentale, con alcuni paesi che mostrano riduzioni nelle emissioni, forse influenzate da cambiamenti economici e politiche ambientali.

Il grafico 3.8 mostra i dati di Oceania e Africa. La scala logaritmica facilita il confronto tra paesi con livelli di sviluppo molto diversi. I dati mostrano un contrasto tra paesi sviluppati con emissioni relativamente elevate e paesi in via di sviluppo con emissioni più basse.

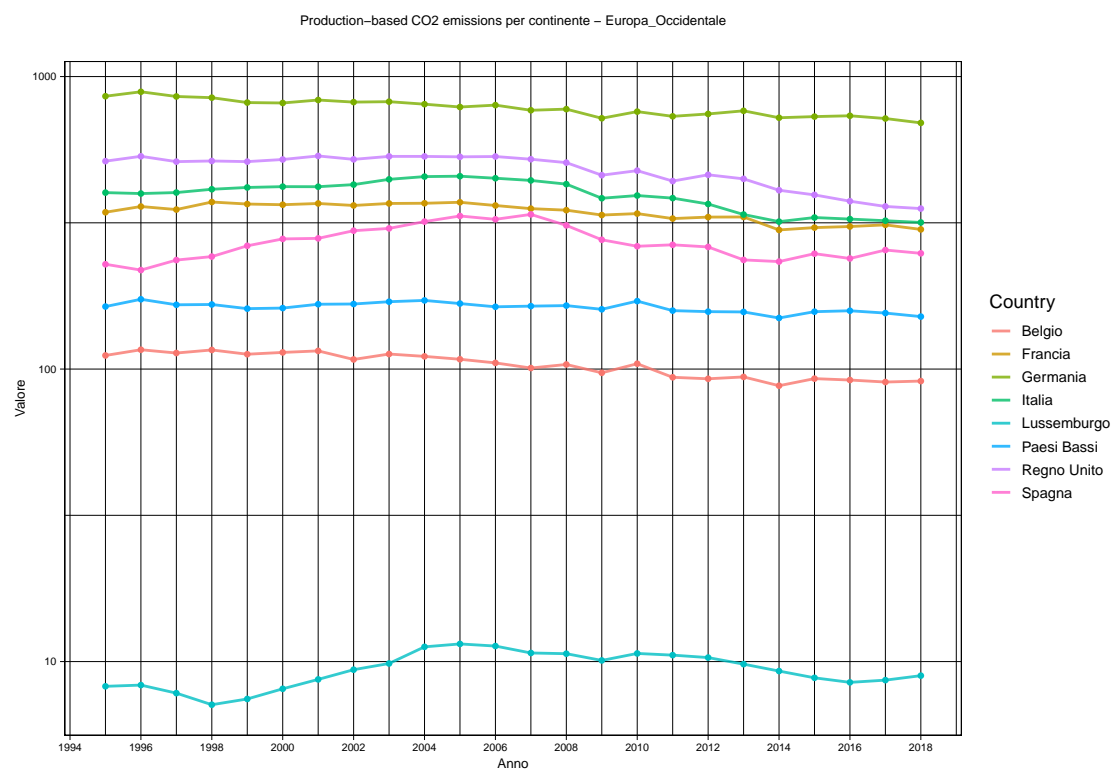


Figura 3.6: Serie storica per i paesi dell'Europa Occidentale.

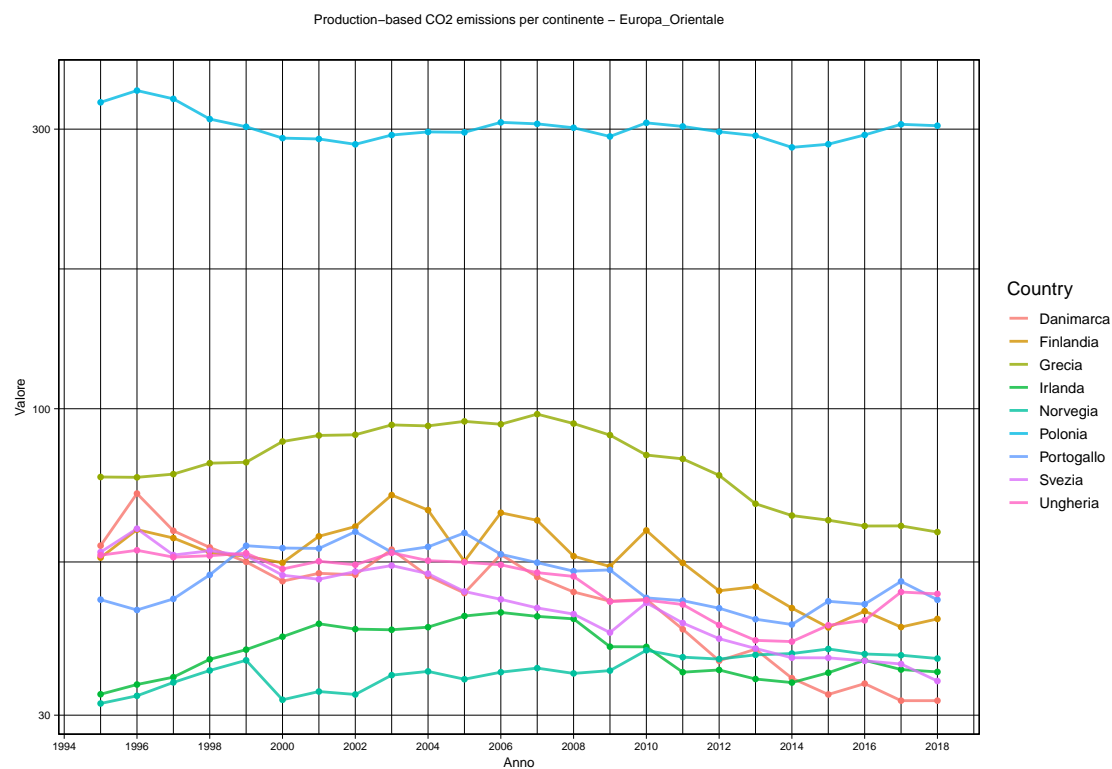


Figura 3.7: Serie storica per i paesi dell'Europa Orientale.

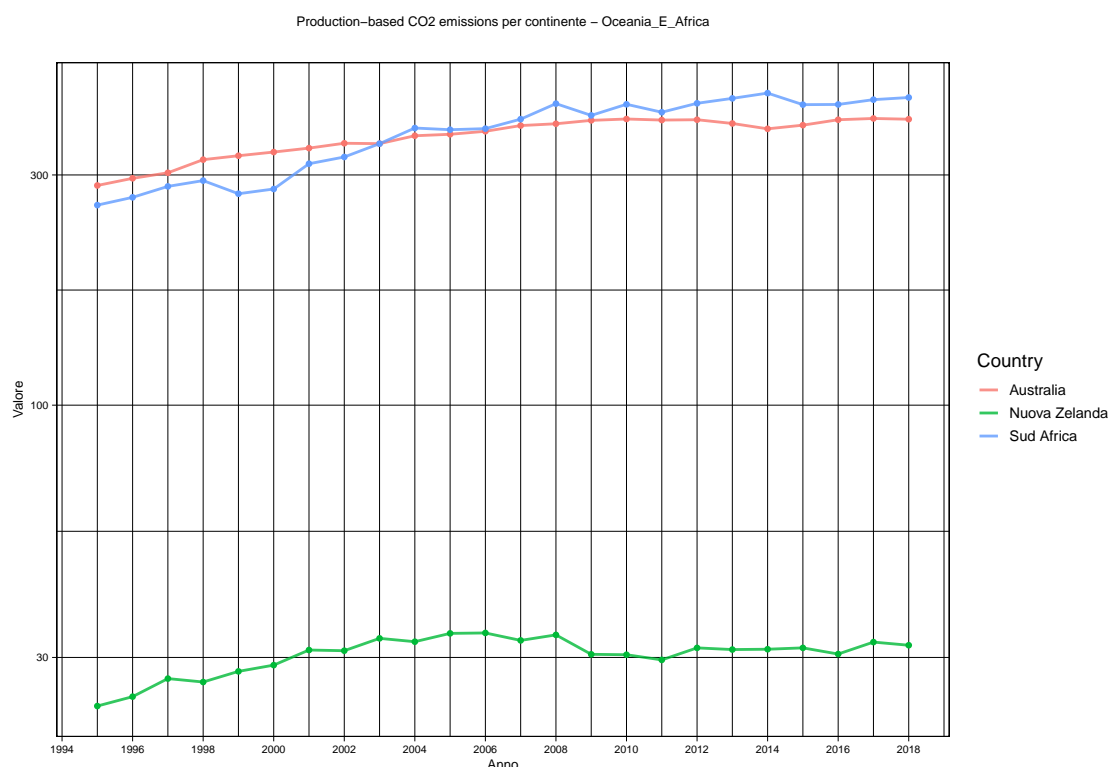


Figura 3.8: Serie storica per i continenti: Oceania e Africa.

3.2.2 Emissioni di CO₂ basate sulla domanda

Le emissioni di CO₂ basate sulla domanda sono espresse in milioni di tonnellate metriche. Le emissioni basate sulla domanda riflettono la CO₂ emessa dall'utilizzo energetico durante le varie fasi di produzione di beni e servizi consumati nella domanda interna finale, indipendentemente da dove si sono verificate le fasi di produzione. Le tendenze nelle emissioni su questa base integrano così le misure più convenzionali basate sulla produzione.

Le stime delle emissioni di CO₂ incorporate nella domanda interna finale sono ottenute dal dataset dell'OCSE sulle emissioni di CO₂ incorporate nel commercio internazionale (edizione 2021), derivato dal database IO dell'OCSE e dalle statistiche sulle emissioni di CO₂ da combustione di combustibili dell'AIE (International Energy Agency (IEA)) (edizione 2021).

Il **Boxplot** 3.9 rappresenta la distribuzione delle emissioni di CO₂ basate sulla domanda nel periodo in esame. La **media** rappresenta il valore medio delle emissioni di CO₂ basate sulla domanda per tutti i paesi considerati. Un valore medio di 720 milioni di tonnellate indica che, in media, i paesi hanno questa quantità di emissioni legate al consumo interno. La **mediana**, essendo significativamente inferiore alla media, indica che più della metà dei paesi ha emissioni di CO₂ basate sulla domanda inferiori a 293 milioni di tonnellate. Questa differenza rispetto alla media suggerisce una **distribuzione asimmetrica**. Il **valore massimo** mostra che il paese con le emissioni più elevate ha raggiunto 9020 milioni di tonnellate, indicando una significativa differenza tra i paesi con le emissioni più basse e più alte. Il **valore minimo** rivela che il paese con le minori emissioni di CO₂ basate sulla domanda ha un livello molto basso, evidenziando una grande variazione nelle emissioni tra diversi paesi. Una **deviazione standard** di 1426 milioni di tonnellate indica una

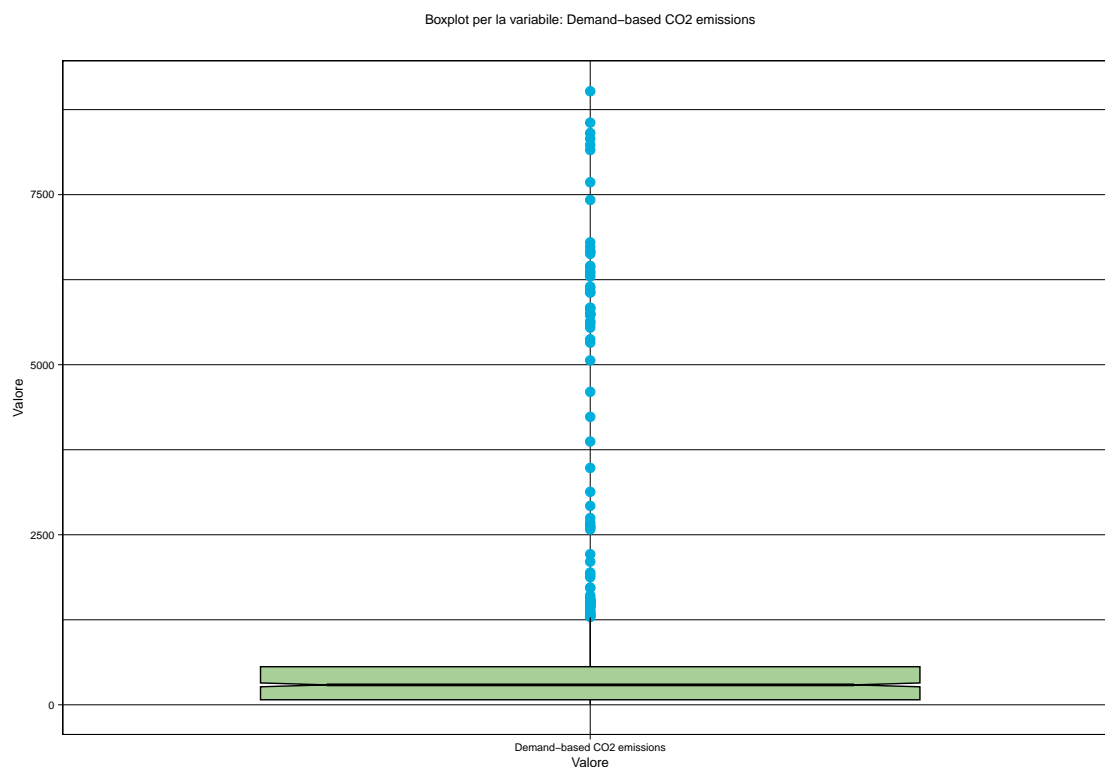
Emissioni di CO ₂ basate sulla domanda	
Media	720
Mediana	293
Simmetria	3.62779
Curtosi	13.16426
Frequenza assoluta	744
Massimo	9020
Minimo	9.59
Deviazione standard	1426
Coefficiente di variazione	1.980383
Varianza	2033552
Q1	74.2
Q3	560

Tabella 3.2: Statistiche descrittive per la variabile Emissioni di CO₂ basate sulla domanda

variazione considerevole nelle emissioni di CO₂ basate sulla domanda tra i paesi, riflettendo diverse pratiche di consumo e livelli di sviluppo.

Il **coefficiente di variazione**, che è il rapporto tra deviazione standard e media, suggerisce una variabilità relativa elevata nelle emissioni di CO₂ tra i paesi.

Il **Primo Quartile (Q1)** e il **Terzo Quartile (Q3)** mostrano che il 50% delle emissioni si colloca in questo intervallo. Un Q1 basso e un Q3 alto indicano che una significativa percentuale di paesi ha emissioni basse, mentre un gruppo più ristretto ha emissioni molto più elevate.

Figura 3.9: Boxplot per la variabile: Emissioni di CO₂ basate sulla domanda

L'**istogramma** 3.10 mostra la frequenza di diversi intervalli di emissioni di CO₂ basate sulla domanda tra i paesi studiati. Ogni barra rappresenta un intervallo di emissioni e la sua altezza indica la frequenza. La maggior parte delle barre è concentrata verso il lato sinistro del grafico, indicando che la maggior parte dei paesi ha livelli relativamente bassi di emissioni basate sulla domanda.

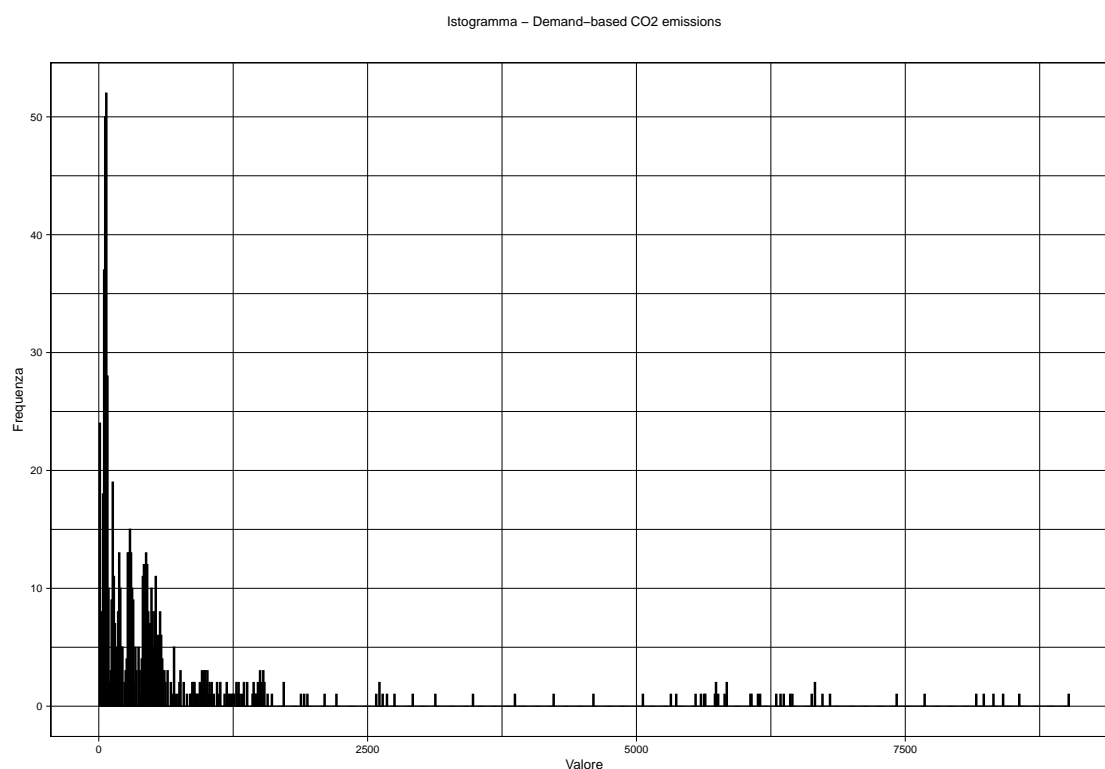


Figura 3.10: Istogramma delle emissioni di CO₂ basate sulla domanda

Il grafico della **funzione di densità** 3.11 rappresenta una stima continua della distribuzione probabilistica delle emissioni di CO₂ basate sulla domanda. La curva liscia indica come si distribuiscono le probabilità per i diversi livelli di emissioni. La curva mostra un picco pronunciato verso il basso valore di emissioni, confermando che un gran numero di paesi ha livelli di emissioni basate sulla domanda relativamente bassi. La coda lunga verso il lato destro del grafico suggerisce la presenza di alcuni paesi con emissioni molto elevate. Questa asimmetria nella distribuzione potrebbe riflettere disuguaglianze economiche e di sviluppo, così come variazioni nelle politiche ambientali e nell'efficienza energetica.

Di seguito, nelle figure 3.12, 3.13, 3.14, 3.15, 3.16 sono riportate le serie storiche per la variabile. Sull'asse orizzontale poniamo gli anni dal 1995 al 2018, mentre sull'asse verticale sono indicate le tonnellate di CO₂ prodotte in **USD/kg** in **scala logaritmica** per rendere il grafico più leggibile, in quanto i paesi che si trovano negli stessi continenti tendono ad avere un andamento simile. All'interno del grafico, ogni linea rappresenta il trend temporale di un singolo paese.

La figura 3.12 mostra le variazioni delle emissioni di CO₂ basate sulla domanda per i paesi del **continente americano**. La scala logaritmica enfatizza le variazioni, soprattutto per paesi con livelli di emissioni diversi. La crisi economica del 2002 in Argentina è evidenziata da un marcato

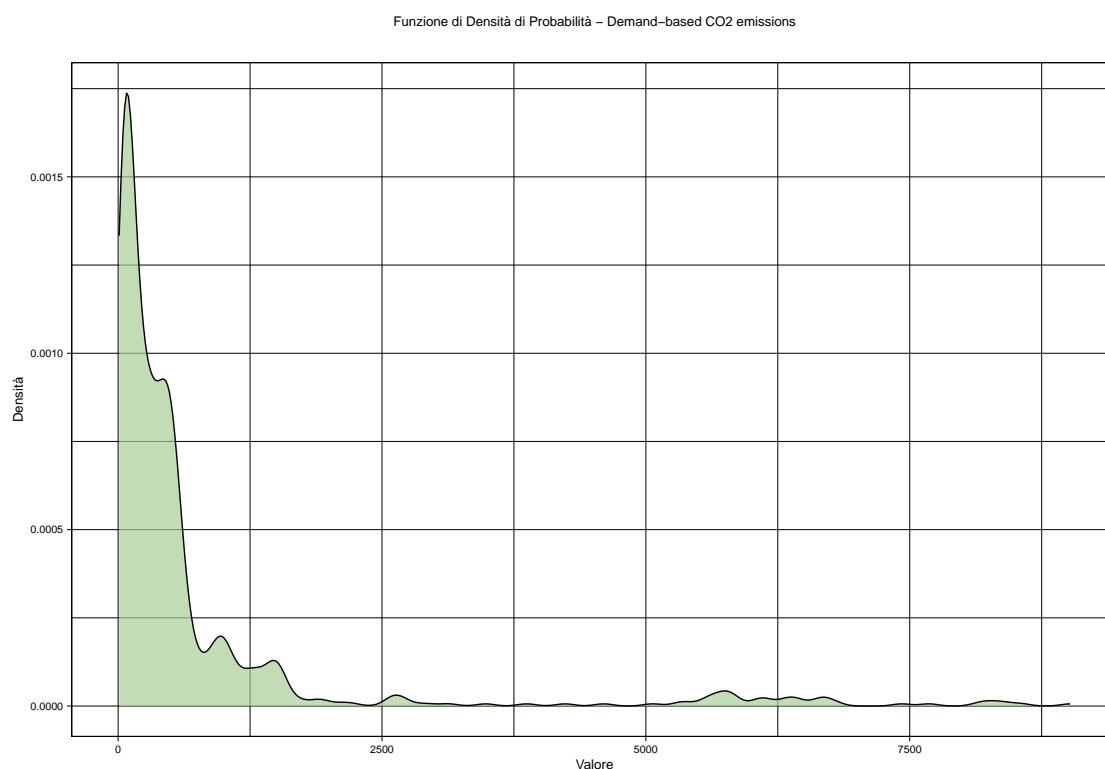


Figura 3.11: Funzione di densità delle emissioni di CO₂ basate sulla domanda

calo delle emissioni, dimostrando come gli eventi economici possano influenzare direttamente l'impatto ambientale. A differenza delle emissioni basate sulla produzione, questo grafico mostra in modo più netto il legame tra l'economia e l'impatto ambientale.

Il grafico 3.13 presenta l'andamento temporale delle emissioni di CO₂ basate sulla domanda per **i paesi asiatici**. La scala logaritmica facilita la comprensione delle tendenze in paesi con livelli di emissione molto diversi. Oltre alla rapida crescita delle emissioni in paesi come Cina, India e Indonesia, il grafico mostra le fluttuazioni economiche in Russia, con una netta riduzione delle emissioni durante la crisi finanziaria degli anni '98-'00, e la crisi del Sud-Est asiatico che ha influenzato le emissioni in Indonesia e Corea del Sud.

La figura 3.14 illustra l'evoluzione delle emissioni di CO₂ basate sulla domanda nei **paesi dell'Europa Occidentale**, offrendo una prospettiva dettagliata sui cambiamenti annuali. Le tendenze generali mostrano un calo delle emissioni in paesi come Grecia, Danimarca, Finlandia e Svezia, riflettendo sia le politiche ambientali avanzate sia l'impatto della recessione economica.

Il grafico 3.15 rappresenta le variazioni delle emissioni di CO₂ basate sulla domanda per **i paesi dell'Europa Orientale**, mettendo in evidenza le differenze tra i paesi della regione. Le variazioni mostrano una maggiore eterogeneità rispetto all'Europa Occidentale, con alcuni paesi che mostrano una riduzione delle emissioni, probabilmente influenzati da cambiamenti economici e politiche ambientali.

La figura 3.16 mostra le emissioni di CO₂ basate sulla domanda per **paesi dell'Oceania e del-**

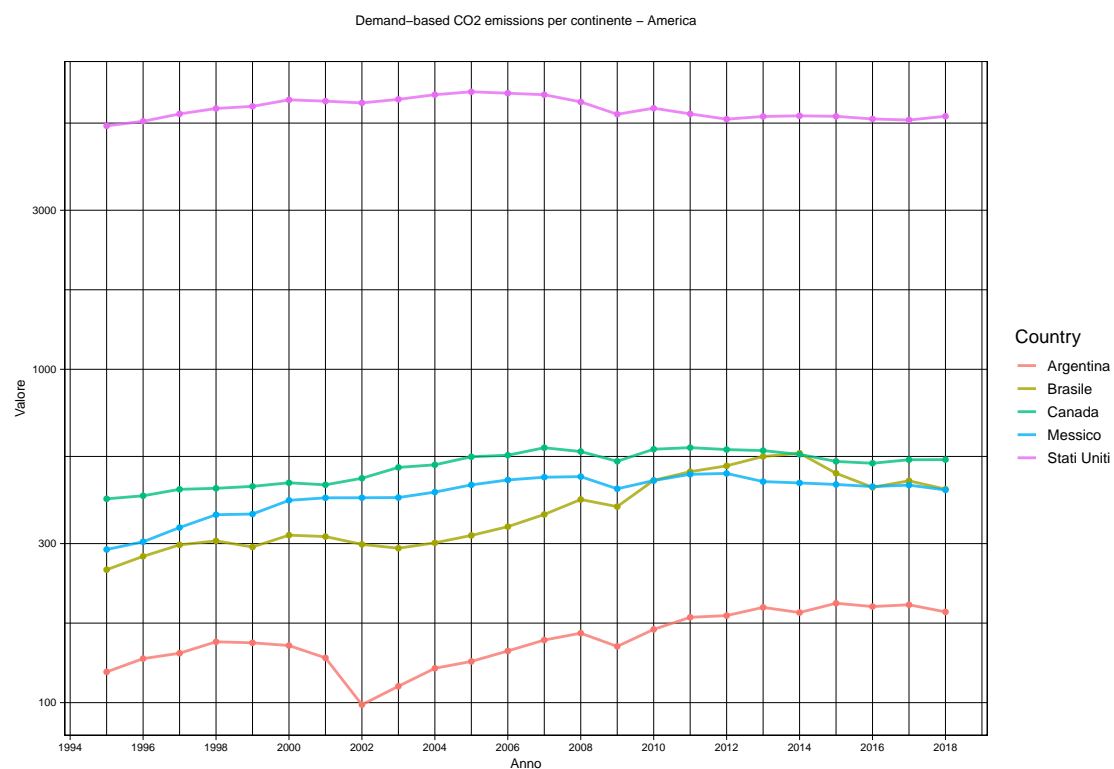
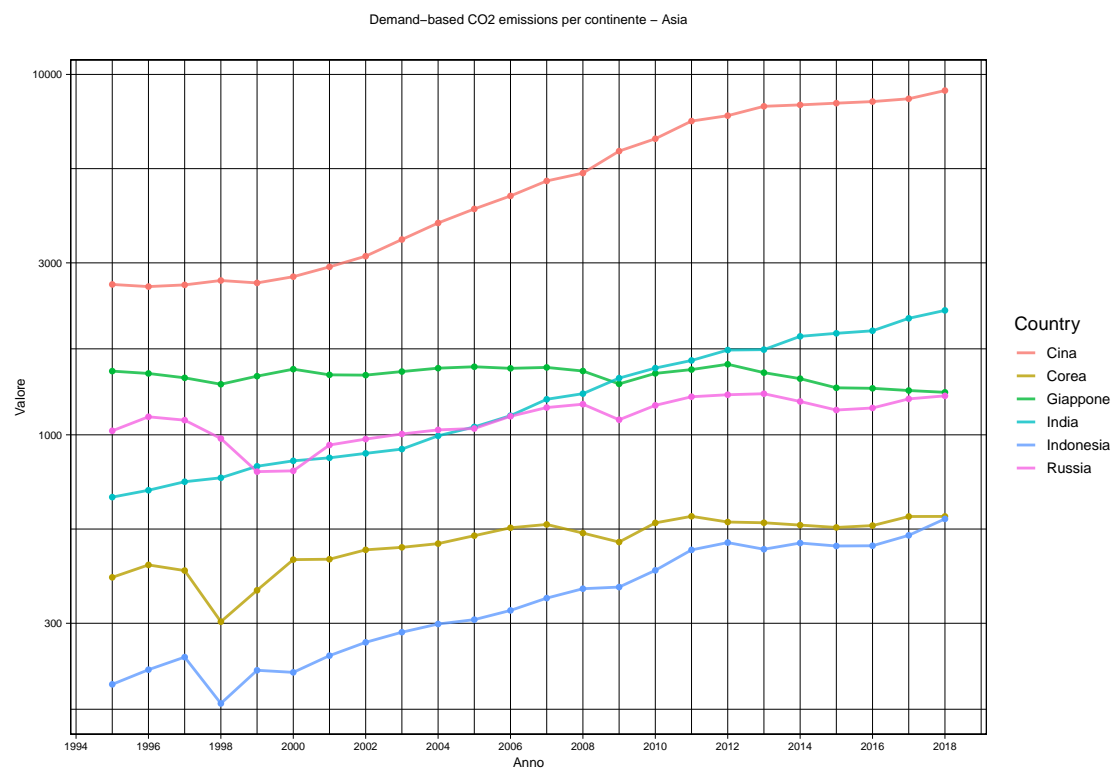
Figura 3.12: Serie storica delle emissioni di CO₂ basate sulla domanda nel continente americano

Figura 3.13: Serie storica per il continente asiatico

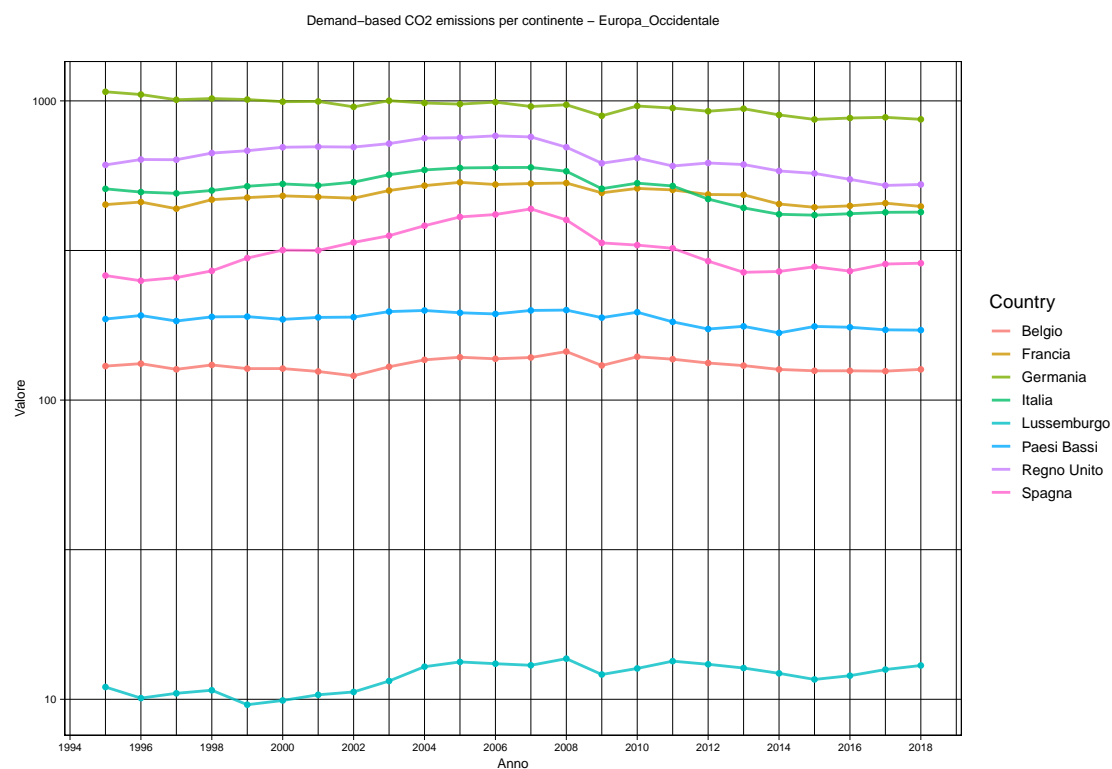


Figura 3.14: Serie storica per i paesi dell'Europa Occidentale

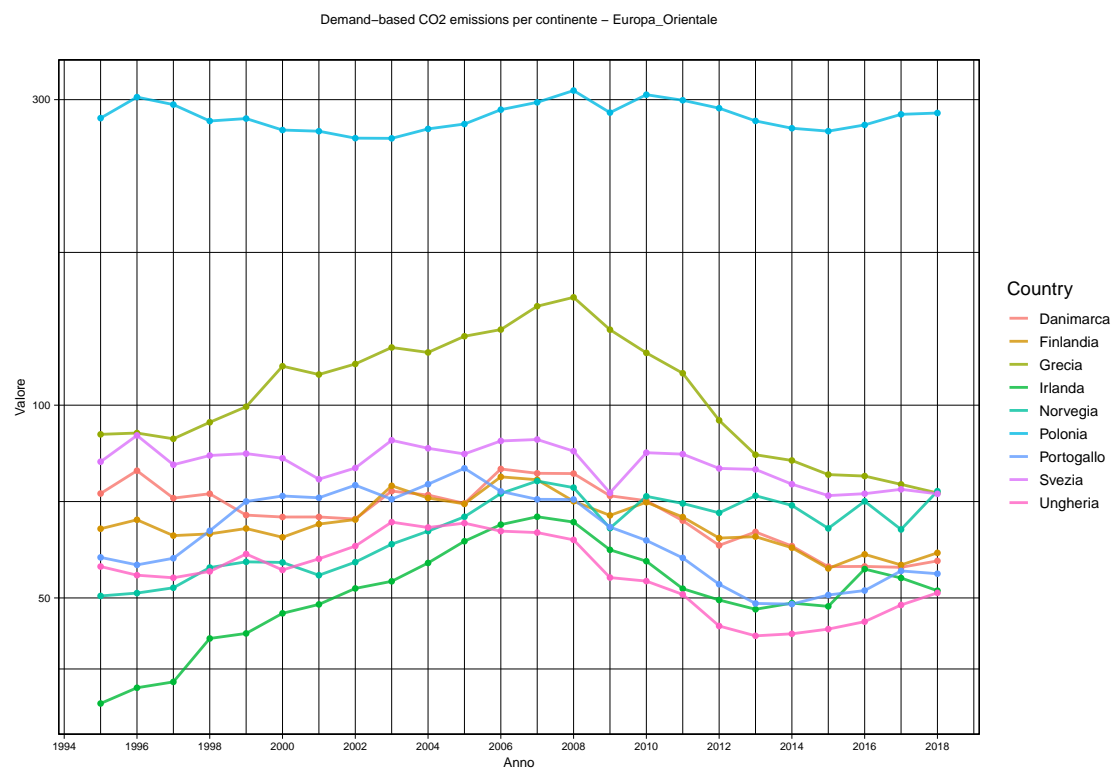


Figura 3.15: Serie storica per i paesi dell'Europa Orientale

l'Africa, confrontando regioni con profili di sviluppo economico e ambientale diversi. Il grafico sottolinea la diversità delle traiettorie di emissione tra i paesi sviluppati e quelli in via di sviluppo, con alcune nazioni che mostrano emissioni relativamente stabili e altre con variazioni significative dovute a cambiamenti economici e sociali.

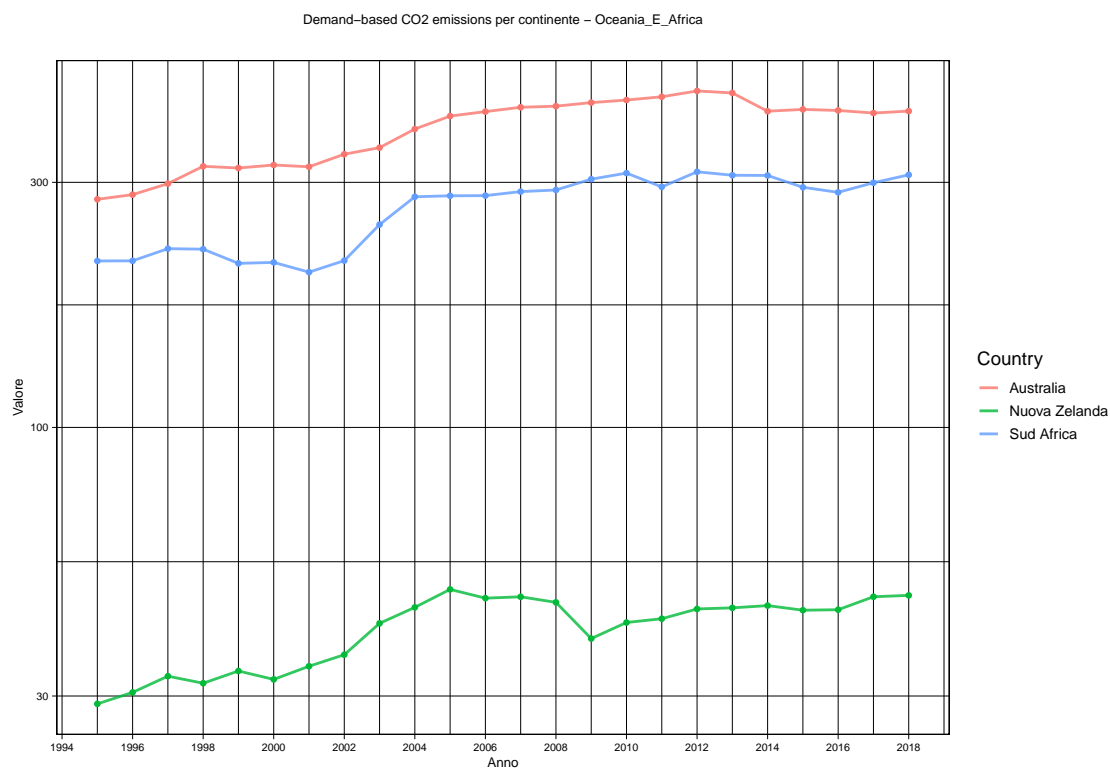


Figura 3.16: Serie storica per i continenti: Oceania e Africa

3.2.3 Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale

Le fonti di acqua potabile migliorate includono: acqua corrente nelle abitazioni, cortili o terreni, rubinetti pubblici o fontanelle, pozzi artesiani o perforati, pozzi scavati protetti, sorgenti protette, acqua in bottiglia, acqua consegnata a domicilio e acqua piovana. Le fonti idriche migliorate dovrebbero essere situate all'interno delle proprietà, disponibili quando necessarie e prive di contaminazione fecale (e dei contaminanti chimici prioritari).

I dati sull'accesso a fonti di acqua potabile migliorate provengono da:

1. **Fondo delle Nazioni Unite per l'Infanzia (UNICEF)/Organizzazione Mondiale della Sanità (OMS) Joint Monitoring Programme for Water Supply and Sanitation (JMP WSS):** Il Joint Monitoring Programme for Water Supply and Sanitation (JMP WSS) è il principale programma di monitoraggio globale per l'acqua e l'igiene. Raccoglie dati annualmente su accesso all'acqua potabile, servizi igienici e igiene di oltre 90 paesi in via di sviluppo.
2. **Organizzazione Mondiale della Sanità (Organizzazione Mondiale della Sanità (OMS)):** L'Organizzazione Mondiale della Sanità (OMS) fornisce dati e linee guida sulla qualità dell'acqua potabile e sui sistemi di approvvigionamento idrico.
3. **Organizzazione delle Nazioni Unite per lo Sviluppo Sostenibile (Organizzazione delle Nazioni Unite per lo Sviluppo Sostenibile (UNDESA)):** L'UNDESA fornisce dati e analisi sulla sostenibilità dello sviluppo, inclusa la fornitura di acqua potabile sicura.

Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale	
Media	87.5
Mediana	96
Simmetria	-1.990271
Curtosi	2.534616
Frequenza assoluta	744
Massimo	100
Minimo	23.5
Deviazione standard	20.7143
Coefficiente di variazione	0.2366661
Varianza	429.0822
Q1	91
Q3	99

Tabella 3.3: Statistiche descrittive per la variabile "Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale"

Il **boxplot** nella figura 3.17 offre una rappresentazione visiva della distribuzione della variabile "Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale" durante il periodo preso in considerazione. Questo grafico fornisce importanti informazioni statistiche sulla distribuzione di questa variabile. La **mediana** delle emissioni è del 96%, indicando che per la maggior parte dei paesi considerati, la popolazione ha accesso a fonti d'acqua migliorate con una percentuale prossima al 96%. La **media**, che è leggermente più bassa (87.5%), suggerisce che ci sono alcune nazioni con valori significativamente inferiori rispetto alla mediana, il che potrebbe essere dovuto a forti disparità tra i paesi.

Il boxplot mostra chiaramente la presenza di alcuni **valori estremi inferiori**, evidenziati come punti al di fuori dei "baffi" del grafico. Ad esempio, il **valore minimo** è solamente 23.5%, il che indica che alcune nazioni hanno un accesso molto limitato a fonti d'acqua migliorate. Al contrario, il **valore massimo** indica che alcuni paesi hanno raggiunto un accesso completo all'acqua potabile per la loro popolazione. Questi paesi sono probabilmente altamente sviluppati con infrastrutture idriche efficaci. Una **deviazione standard** relativamente alta rispetto alla media indica un'ampia varietà nei livelli di accesso all'acqua potabile tra i paesi. Il **coefficiente di variazione** relativamente basso suggerisce che, nonostante alcune variazioni, la maggior parte dei paesi si avvicina alla media. I **quartili Q1 (91%) e Q3 (99%)** mostrano che il 50% dei paesi ha un livello di accesso all'acqua potabile tra il 91% e il 99%, indicando che la maggior parte dei paesi ha un buon accesso.

La **simmetria negativa** della distribuzione, indicata da un valore di skewness pari a -1.990271, conferma la presenza di una coda lunga verso valori più bassi. Ciò significa che alcune nazioni hanno un accesso notevolmente inferiore a fonti d'acqua migliorate, influenzando in maniera significativa la distribuzione complessiva. Inoltre, la **curtosi** della distribuzione è di 2.534616, indicando una certa "pesantezza" delle code della distribuzione, il che suggerisce una maggiore variabilità rispetto a una distribuzione normale.

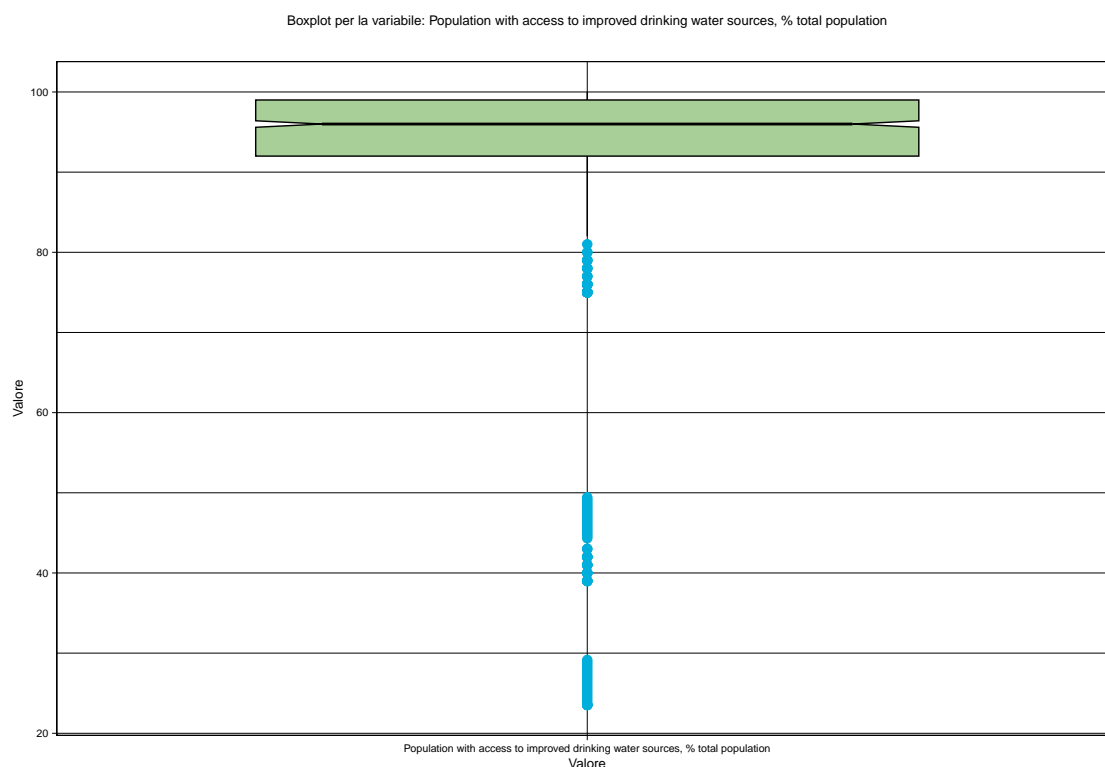


Figura 3.17: Boxplot per la variabile: Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale

In figura 3.18 è mostrato l'istogramma che rappresenta la distribuzione della variabile. Sull'asse delle ascisse sono presenti i possibili valori assunti, mentre sull'asse delle ordinate la frequenza di ciascun valore. Ad esempio, ci sono molti paesi con valori vicini al 100%, il che indica che la

maggior parte di questi paesi ha un alto accesso a fonti d'acqua migliorate. L'istogramma mostra un' **asimmetria negativa**, confermata dal valore negativo di skewness. Questa distribuzione può avere diverse interpretazioni: potrebbe significare che la maggior parte dei paesi sta facendo progressi nel migliorare l'accesso all'acqua potabile, ma ci sono ancora alcune nazioni che affrontano sfide significative in questo settore.

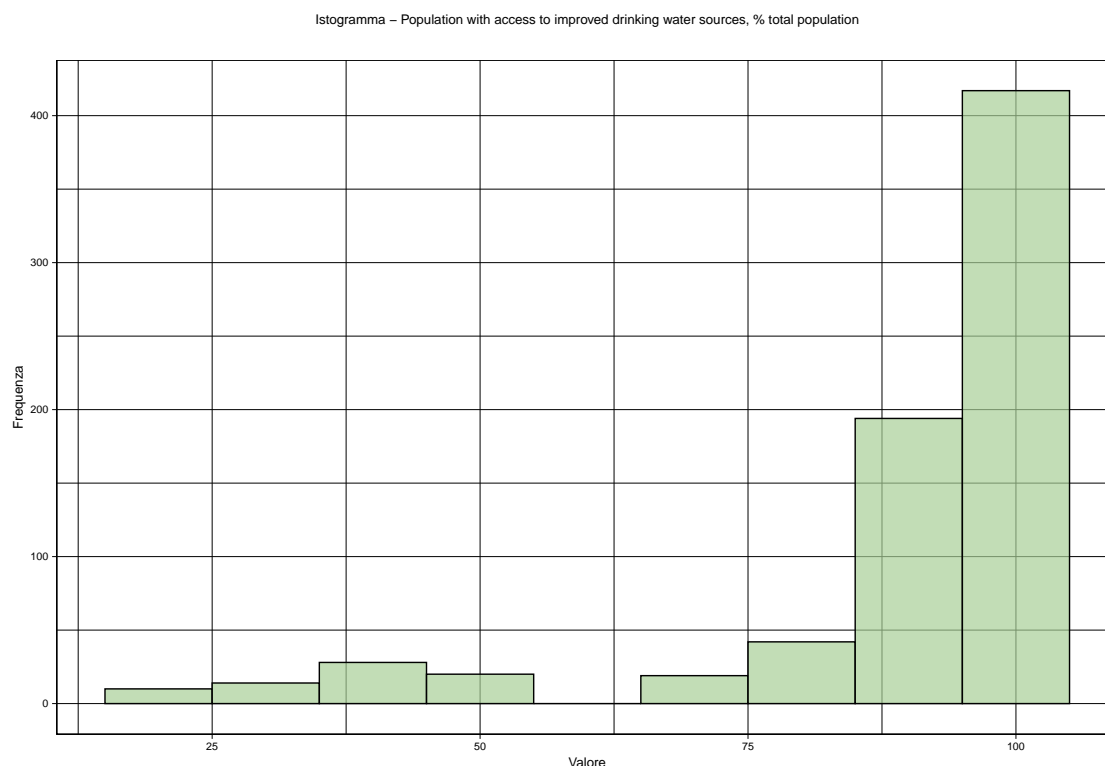


Figura 3.18: Istogramma per la variabile: Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale

La figura 3.19 mostra la **funzione di densità** per la variabile. La curva rappresenta la distribuzione continua della probabilità per i diversi livelli di accesso all'acqua potabile tra i paesi. La curva mostra un picco pronunciato verso il lato destro, indicando che un gran numero di paesi ha un alto livello di accesso all'acqua potabile. Tuttavia, si nota anche una **coda lunga** verso il lato sinistro del grafico. Questa coda rappresenta i paesi con livelli di accesso all'acqua significativamente più bassi, confermando la presenza di asimmetria nella distribuzione. La forma asimmetrica della curva, con una coda verso i valori più bassi, corrisponde al valore negativo di skewness nella tabella statistica. Questo suggerisce che, sebbene la maggior parte dei paesi abbia un buon accesso all'acqua, ci sono alcune nazioni con situazioni critiche.

Di seguito, nelle figure 3.20, 3.21, 3.22, 3.23, 3.24, sono riportate le serie storiche per la variabile. Sull'asse orizzontale sono indicati gli anni dal 1995 al 2018, mentre sull'asse verticale è riportata la percentuale della popolazione con accesso a fonti d'acqua potabile. All'interno del grafico, ogni linea rappresenta il trend temporale di un singolo paese.

Per il continente **Americano** (mostrato in figura 3.20) si nota che i paesi sviluppati come Canada

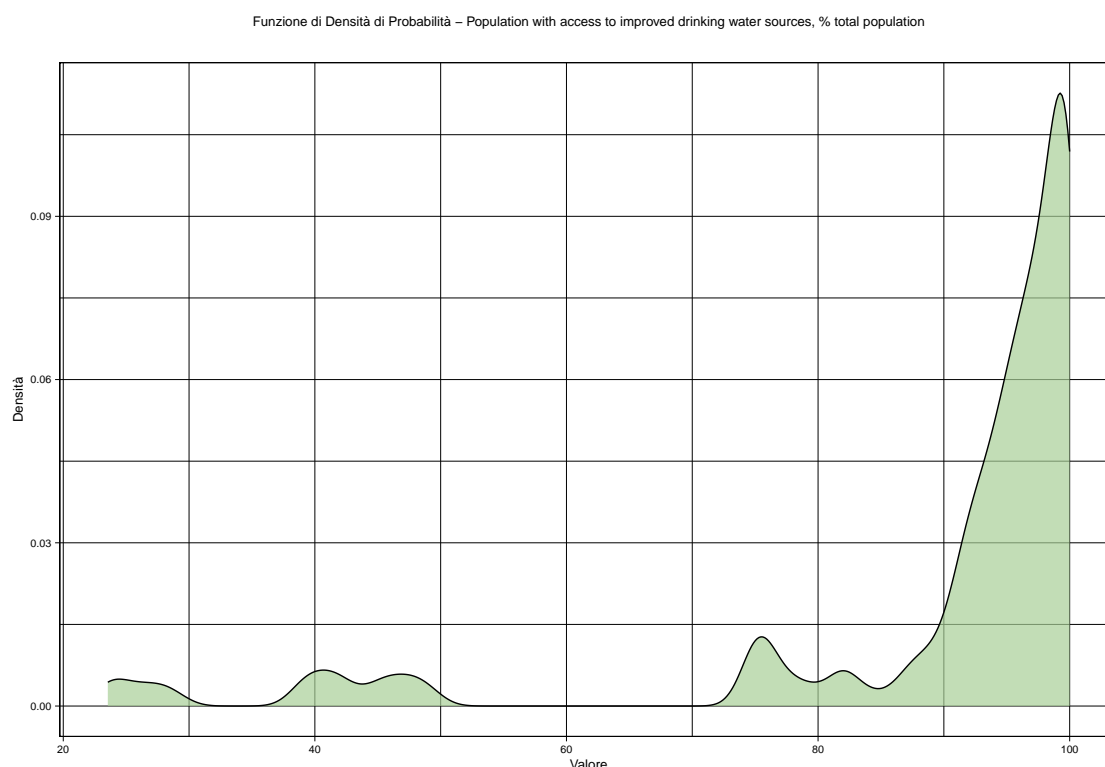


Figura 3.19: Funzione di densità per la variabile: Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale

e Stati Uniti hanno un andamento quasi costante prossimo al 100%. Nei paesi in via di sviluppo, come Argentina e Brasile, si osserva un miglioramento progressivo della percentuale di accesso all'acqua. Al contrario, il Messico, un paese con maggiori sfide economiche, presenta percentuali inferiori.

Per il continente **Asiatico** (mostrato in figura 3.21) si nota che anche in questo caso i paesi sviluppati come Giappone e Corea del Sud hanno un andamento quasi costante prossimo al 100%. Al contrario, Cina e Indonesia mostrano percentuali molto più basse, indicando significative sfide nell'accesso all'acqua potabile.

I paesi dell'**Europa Occidentale** (figura 3.22) mostrano livelli molto elevati e stabili di accesso all'acqua potabile, grazie a infrastrutture ben sviluppate e politiche efficaci.

Similmente all'Europa Occidentale, l'**Europa Orientale** (figura 3.23) mostra un alto livello di accesso, con una tendenza costante e stabile.

La situazione in **Oceania e Africa** è più variegata (figura 3.24). Mentre i paesi oceanici sviluppati hanno un accesso quasi universale, il Sud Africa mostra un progresso nel tempo, riflettendo disparità economiche e sfide infrastrutturali.

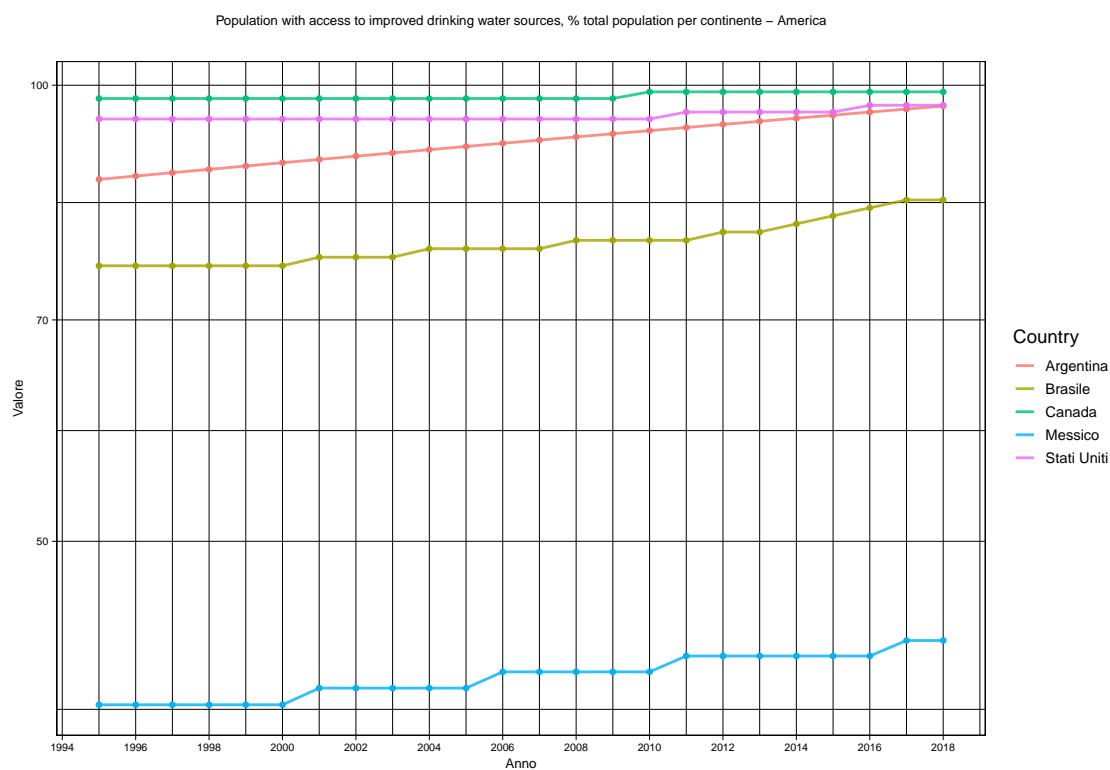


Figura 3.20: Serie storica per il continente Americano: Popolazione con accesso a fonti di acqua potabile migliorate

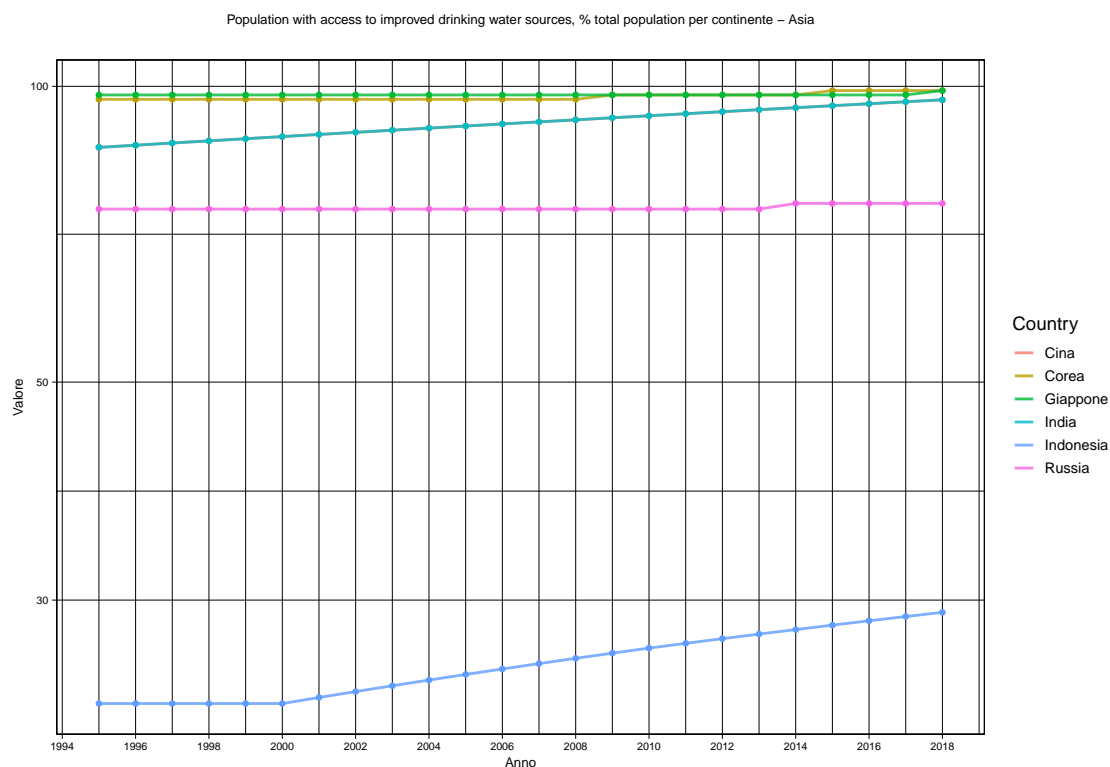


Figura 3.21: Serie storica per il continente Asiatico: Popolazione con accesso a fonti di acqua potabile migliorate

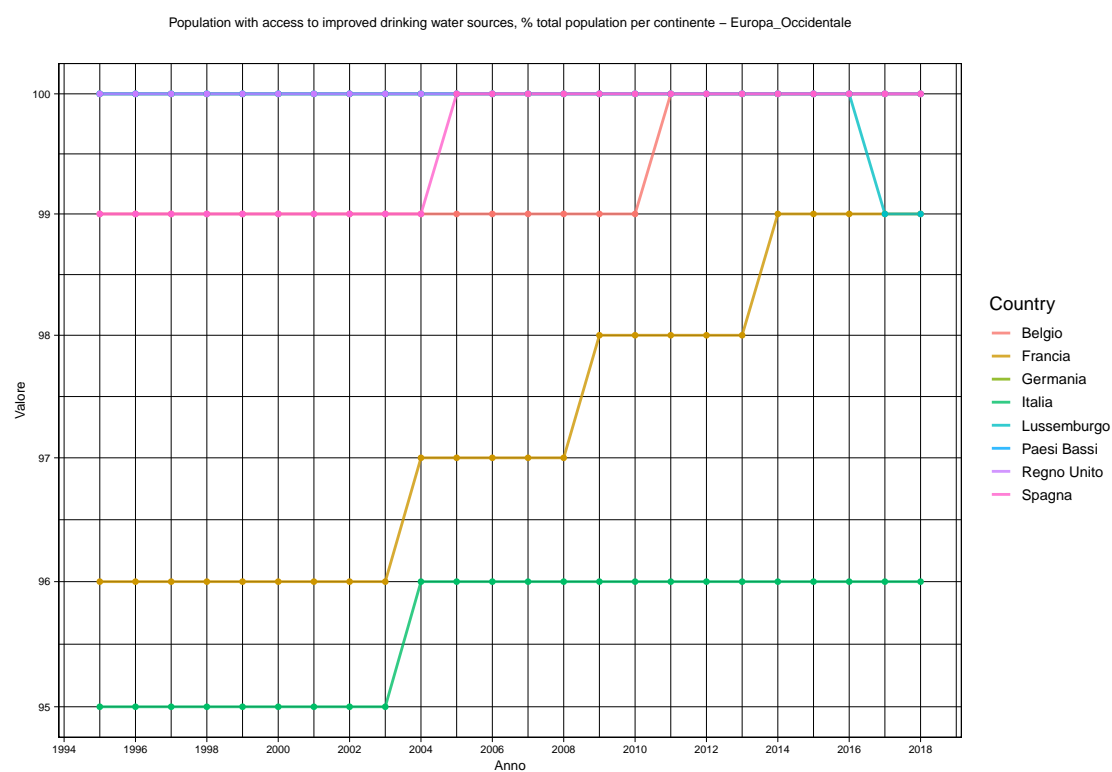


Figura 3.22: Serie storica per i paesi dell'Europa Occidentale: Popolazione con accesso a fonti di acqua potabile migliorate

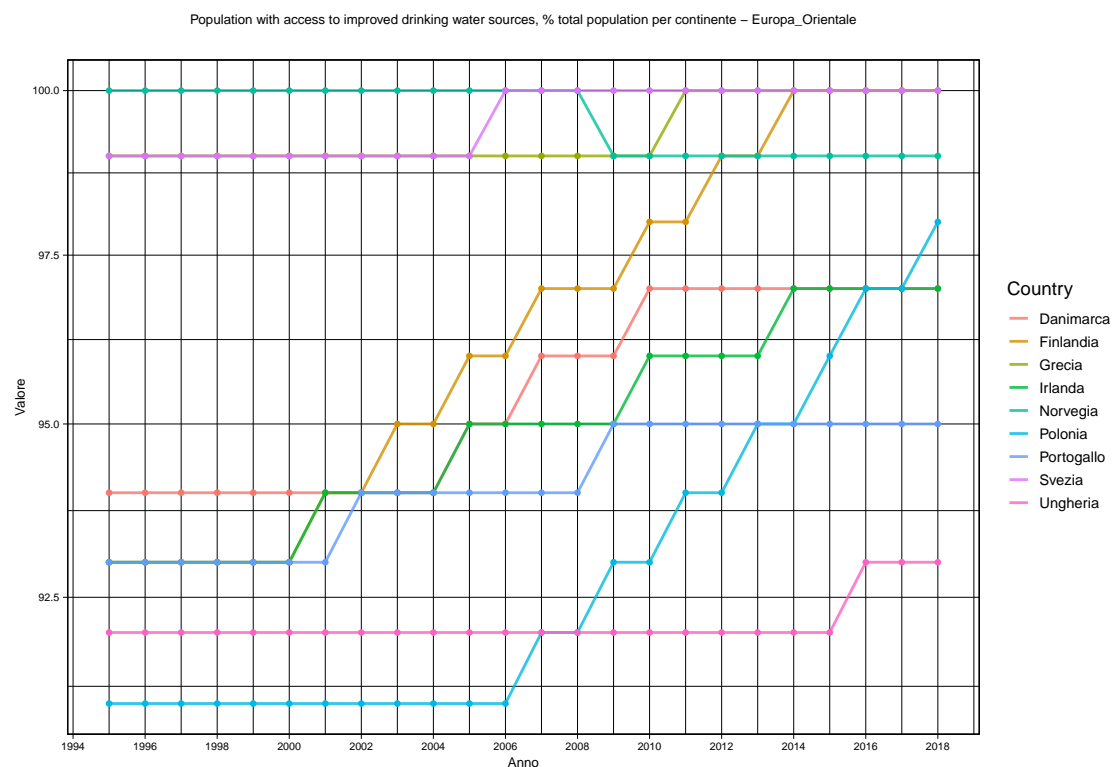


Figura 3.23: Serie storica per i paesi dell'Europa Orientale: Popolazione con accesso a fonti di acqua potabile migliorate

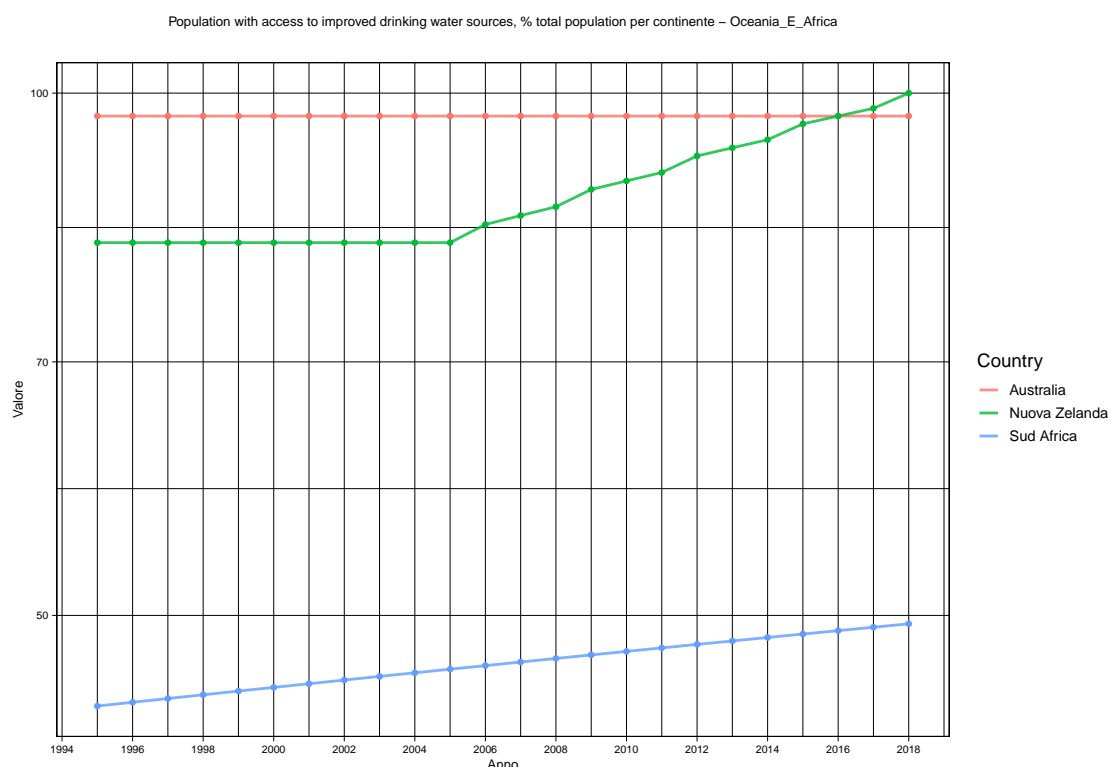


Figura 3.24: Serie storica per i continenti: Oceania e Africa: Popolazione con accesso a fonti di acqua potabile migliorate

3.2.4 Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

Lo Stress Idrico (SI) dalle risorse disponibili misura le estrazioni totali lorde di acqua dolce espresse come percentuale delle risorse totali di acqua dolce rinnovabile disponibile (compresi gli afflussi dai paesi vicini).

Lo SI può essere classificato come:

- **Basso (meno del 10%):** generalmente nessuna grande pressione sulle risorse disponibili.
- **Moderato (10-20%):** la disponibilità di acqua sta diventando un vincolo per lo sviluppo, rendendo necessari investimenti significativi per fornire approvvigionamenti adeguati.
- **Medio-alto (20-40%):** implica la gestione sia dell'offerta che della domanda e la necessità di risolvere i conflitti tra gli usi concorrenziali dell'acqua.
- **Alto (oltre il 40%):** indica una grave scarsità e generalmente mostra un utilizzo dell'acqua non sostenibile, che può diventare un fattore limitante per lo sviluppo sociale ed economico.

I dati sulle estrazioni di acqua dolce sono ottenuti dal dataset Freshwater Abstractions (FWA) del database ambientale dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE).

La **media** dello stress idrico è del 25.4%, che rappresenta la percentuale media di prelievo di acqua dolce rispetto alle risorse rinnovabili totali disponibili nei paesi considerati. La **mediana**, che si

Stress idrico, prelievo totale di acqua dolce	
Media	25.4
Mediana	20.1
Simmetria	1.091017
Curtosi	0.6216051
Frequenza assoluta	744
Massimo	101
Minimo	1.46
Deviazione standard	22.19914
Coefficiente di variazione	0.8740704
Varianza	492.8019
Q1	5.45
Q3	39.5

Tabella 3.4: Statistiche descrittive per la variabile: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

attesta al 20.1%, indica che il 50% dei paesi ha uno stress idrico inferiore a questo valore. La mediana funge da punto di riferimento centrale per valutare la situazione di stress idrico nei paesi.

L'indicatore presenta una leggera **asimmetria positiva** nella sua distribuzione, come indicato dal valore di **skewness** pari a 1.091017. Questa asimmetria suggerisce che alcuni paesi hanno livelli di stress idrico notevolmente superiori alla media. La **curtosi**, che ammonta a 0.6216051, indica che la distribuzione dell'indicatore è leggermente appiattita rispetto a una distribuzione normale. Questo suggerisce che la variabilità nei livelli di stress idrico potrebbe essere meno accentuata rispetto a una distribuzione più "appuntita".

L'indicatore registra valori estremi, con un **massimo** del 101%, il che indica che alcuni paesi superano la disponibilità rinnovabile di acqua dolce, situazione altamente insostenibile. D'altra parte, il **valore minimo** dell'1.46% rappresenta uno stress idrico relativamente basso. La **deviazione standard** è di 22.19914, evidenziando una certa variabilità nei livelli di stress idrico tra i paesi considerati. Il **coefficiente di variazione** è dello 0.8740704, rappresentando la variabilità relativa rispetto alla media. Questo valore suggerisce che la variabilità nei livelli di stress idrico è significativa rispetto alla media. La **varianza** ammonta a 492.8019, confermando la presenza di variabilità nei dati.

Il **Primo Quartile (Q1)** è pari al 5.45%, mentre il **Terzo Quartile (Q3)** si attesta al 39.5%. Questi quartili indicano la distribuzione dei dati e suggeriscono che molti paesi affrontano uno stress idrico moderato o alto.

L'**istogramma** in figura 3.26 rappresenta graficamente la distribuzione dei livelli di stress idrico nei paesi considerati. Sull'asse delle ascisse sono riportati i possibili valori assunti dall'indicatore, che rappresentano la percentuale di prelievo di acqua dolce rispetto alle risorse rinnovabili totali disponibili. Sull'asse delle ordinate è riportata la frequenza con cui ciascun valore appare nei dati.

L'istogramma mostra una **distribuzione asimmetrica positiva**, il che significa che la maggior parte dei paesi ha livelli di stress idrico inferiori, concentrati verso sinistra nell'istogramma. Questo è rappresentato da un picco nell'istogramma a valori bassi di stress idrico. Tuttavia, l'istogramma rivela anche la presenza di alcuni paesi con livelli di stress idrico notevolmente più elevati rispetto

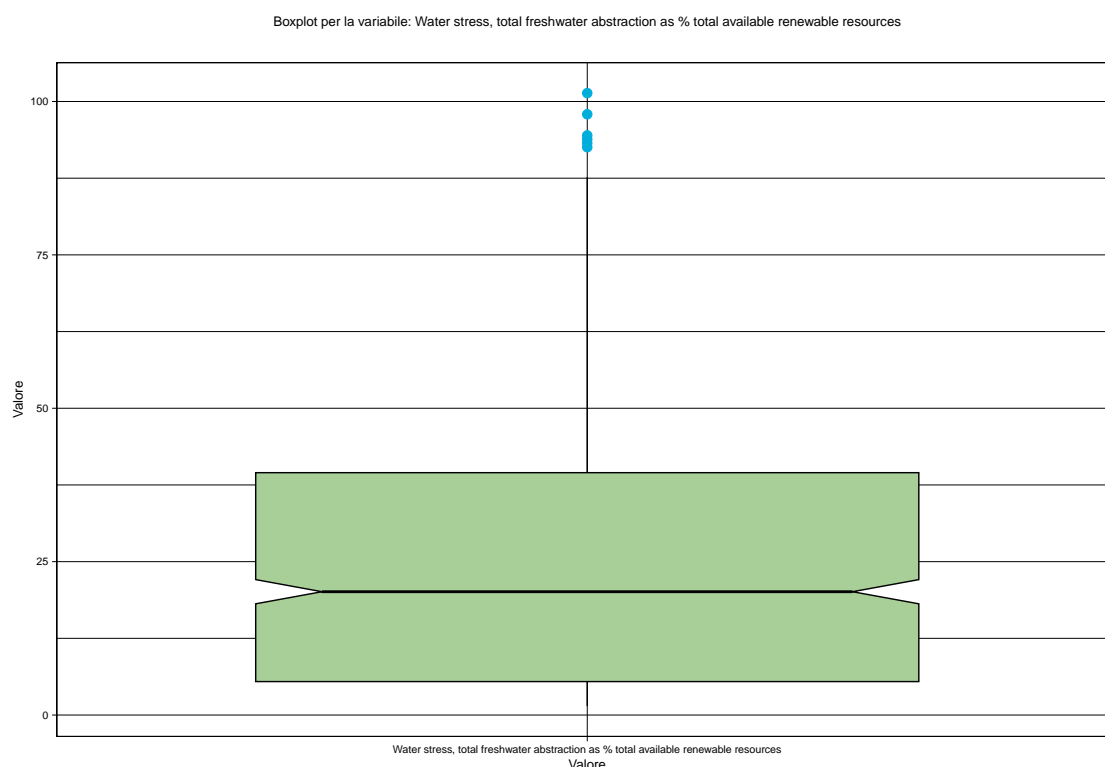


Figura 3.25: Boxplot per la variabile: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

alla media. Questi paesi sono rappresentati da code lunghe verso destra nell'istogramma, indicando che ci sono alcune regioni con una grave scarsità di risorse idriche. Inoltre, è possibile notare che la maggior parte dei paesi si concentra nell'intervallo tra il 0% e il 50% di prelievo di acqua dolce rispetto alle risorse rinnovabili totali disponibili. Questo intervallo rappresenta una gamma di stress idrico da moderato a medio-alto, suggerendo la necessità di gestire sia l'offerta che la domanda di acqua.

La figura 3.27 mostra la **funzione di densità** per lo stress idrico, che rappresenta il prelievo totale di acqua dolce come percentuale delle risorse idriche rinnovabili totali disponibili. La curva lascia della funzione di densità offre una rappresentazione continua della probabilità di diversi livelli di stress idrico nei paesi analizzati. La curva mostra un picco pronunciato per valori bassi di stress idrico, suggerendo che la maggior parte dei paesi ha un livello moderato di prelievo di risorse idriche rispetto alla disponibilità totale. La curva presenta anche una coda lunga verso valori più alti, indicando la presenza di paesi con livelli elevati di stress idrico. Questo può riflettere situazioni di scarsità d'acqua, gestione inefficace delle risorse idriche, o una combinazione di alta domanda e limitata disponibilità di acqua dolce. La distribuzione evidenziata dalla funzione di densità indica che, mentre molti paesi mantengono un equilibrio sostenibile tra prelievo e rinnovabilità delle risorse idriche, esiste un gruppo significativo di paesi che affronta una grave pressione sulle loro risorse idriche. Questo richiede politiche mirate per una gestione più efficiente e sostenibile dell'acqua, specialmente in quelle regioni che mostrano un alto livello di stress idrico.

Di seguito, nelle figure 3.28, 3.29, 3.30, 3.31, 3.32, sono riportate le serie storiche per la variabile.

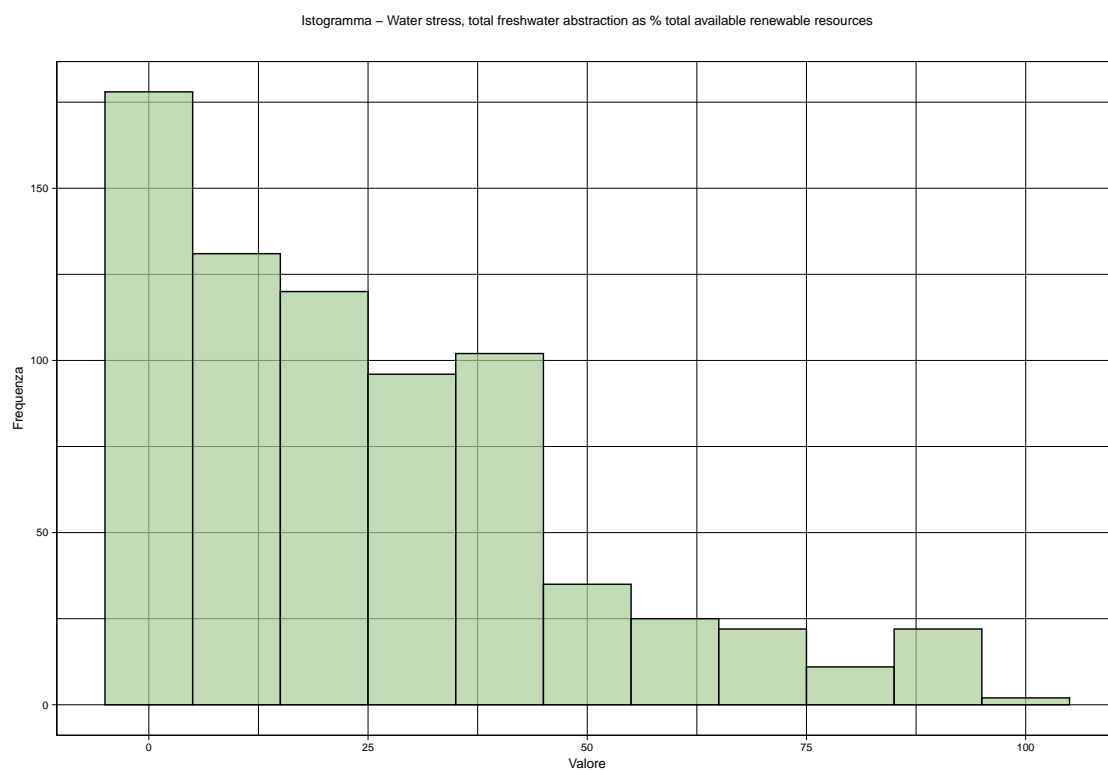


Figura 3.26: Istogramma per la variabile: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

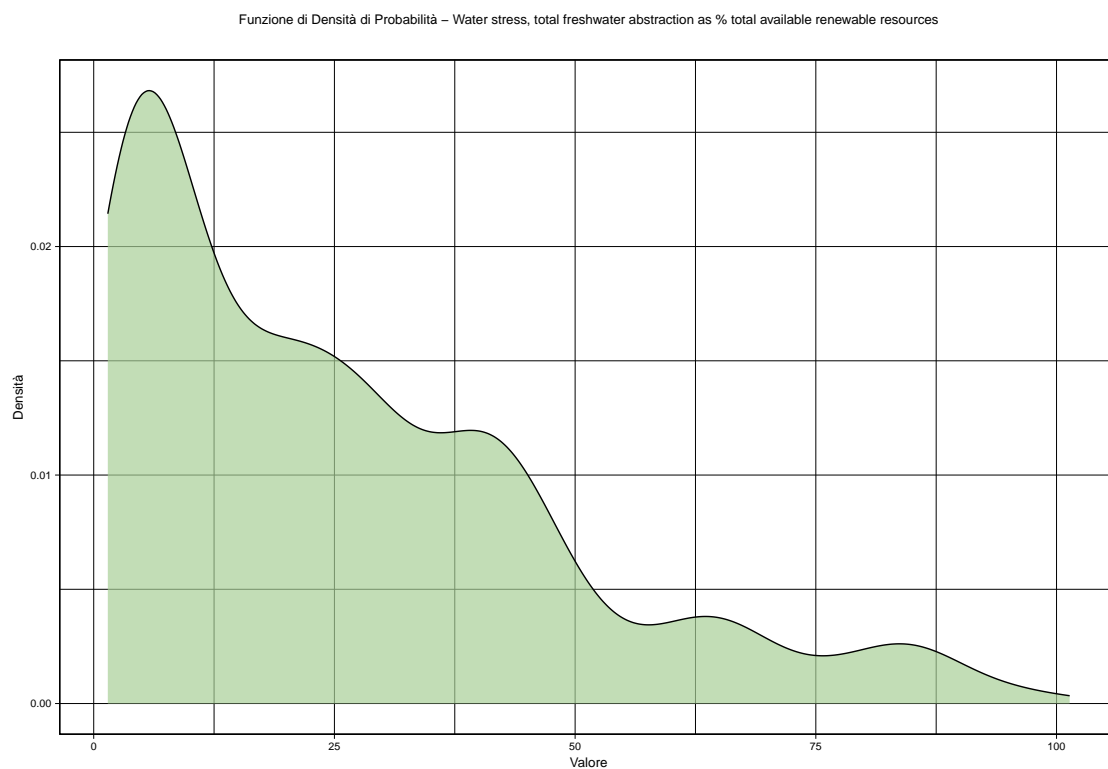


Figura 3.27: Funzione di densità per la variabile: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

Sull'asse orizzontale sono indicati gli anni dal 1995 al 2018, mentre sull'asse verticale è riportata la percentuale di stress idrico. All'interno del grafico, ogni linea rappresenta il trend temporale di un singolo paese.

Nel continente **Americano** (mostrato in figura 3.28), si osservano diverse tendenze nei paesi considerati. Il Messico ha mostrato una rapida crescita del livello di stress idrico nel periodo considerato, mentre gli Stati Uniti hanno sperimentato una diminuzione del livello di stress idrico. Al contrario, Argentina, Brasile e Canada hanno mantenuto livelli bassi di pressione sulle risorse idriche.

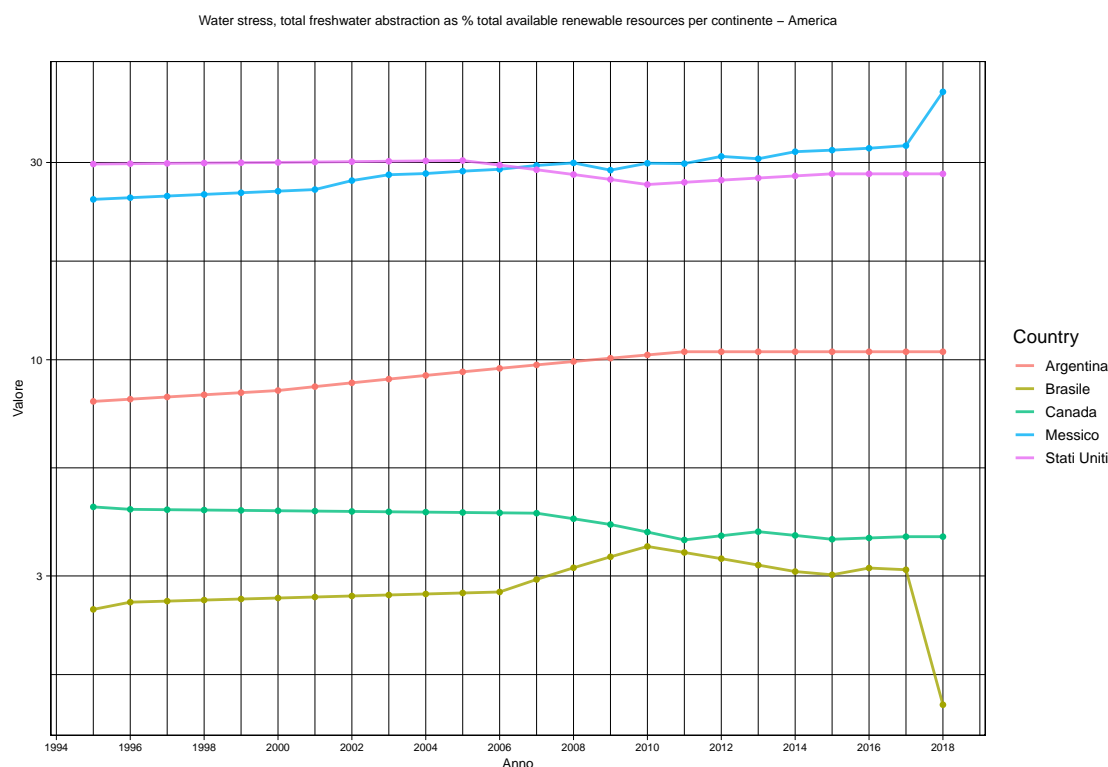


Figura 3.28: Serie storica per il continente Americano: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

In **Asia** (figura 3.29), la situazione è molto eterogenea. Paesi come l'India e la Corea del Sud mostrano livelli elevati di stress idrico, a volte superando il 40%. Questo può essere attribuito a una combinazione di alta densità di popolazione, rapido sviluppo industriale e agricoltura intensiva.

Nel continente **Europeo**, sia **Occidentale** (figura 3.30) che **Orientale** (figura 3.31), si osserva una situazione molto variegata. Pur avendo infrastrutture avanzate, alcuni paesi affrontano sfide crescenti, potenzialmente legate ai cambiamenti climatici e all'incremento dell'uso industriale e agricolo dell'acqua. Paesi come il Belgio hanno raggiunto il 101% di utilizzo delle risorse idriche ma poi hanno mostrato un trend in diminuzione, arrivando al 50% alla fine delle serie storiche. Un trend simile in diminuzione si osserva anche per la Germania. Al contrario, i Paesi Bassi e la Danimarca mostrano un aumento dello stress idrico nel periodo considerato. Gli altri paesi hanno seguito trend più regolari.

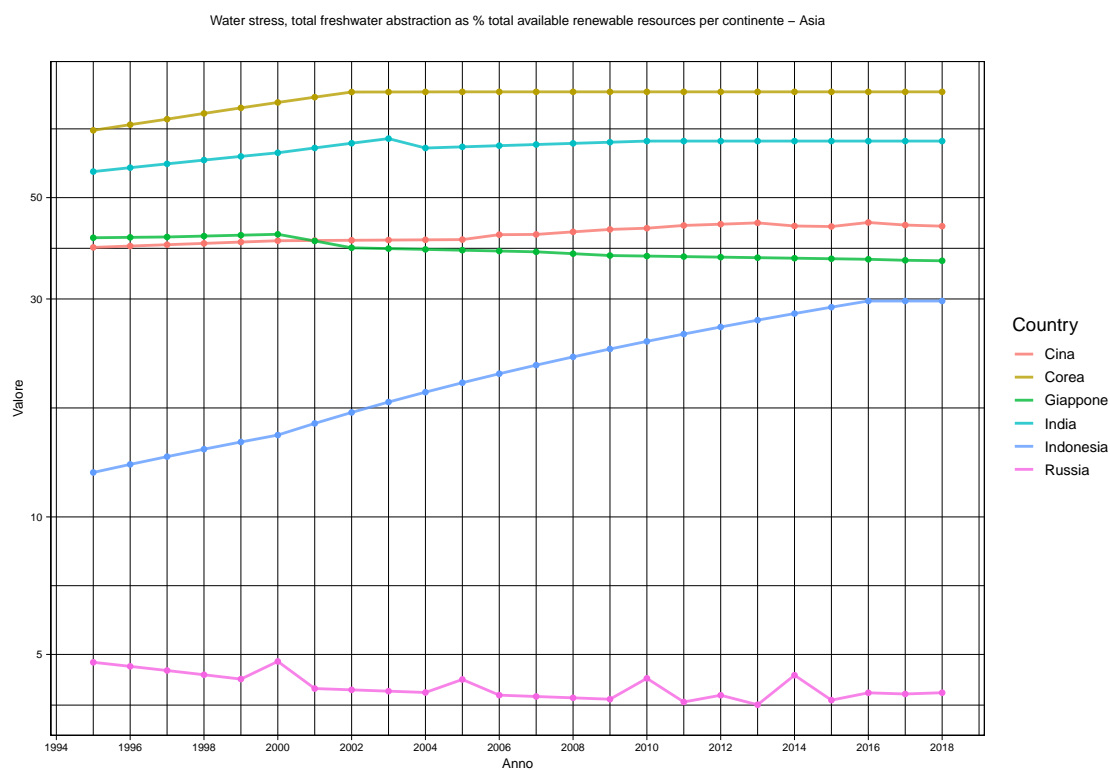


Figura 3.29: Serie storica per il continente Asiatico: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

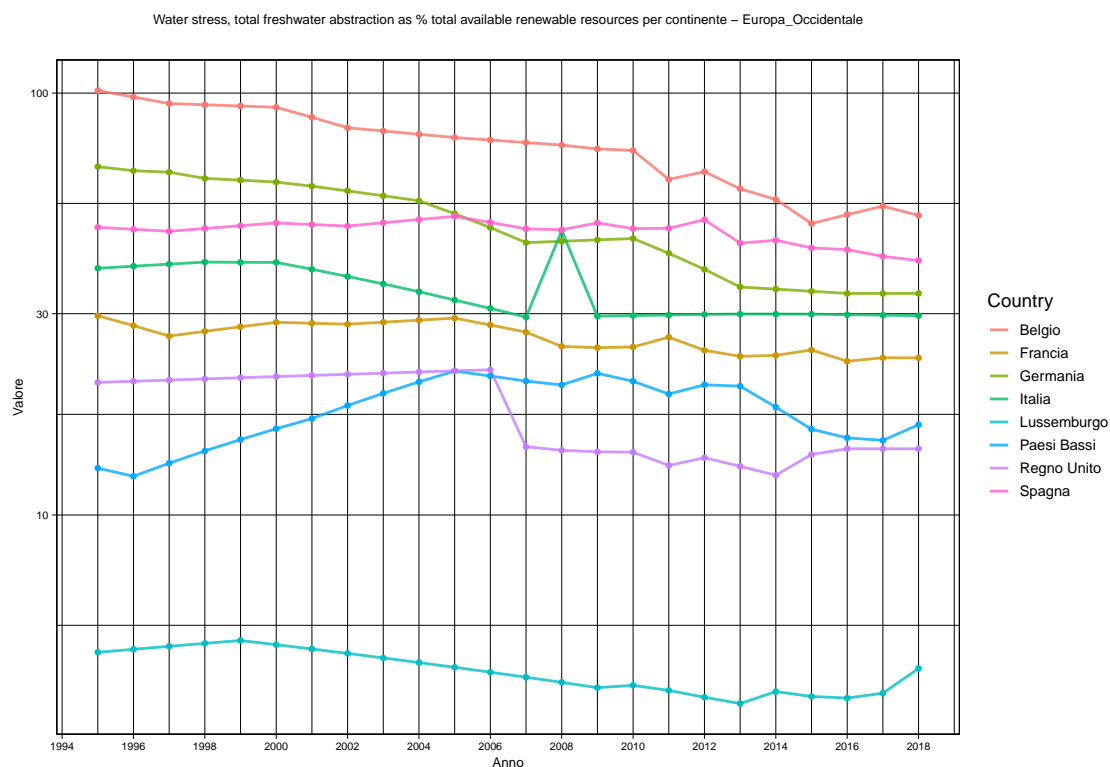


Figura 3.30: Serie storica per i paesi dell'Europa Occidentale: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

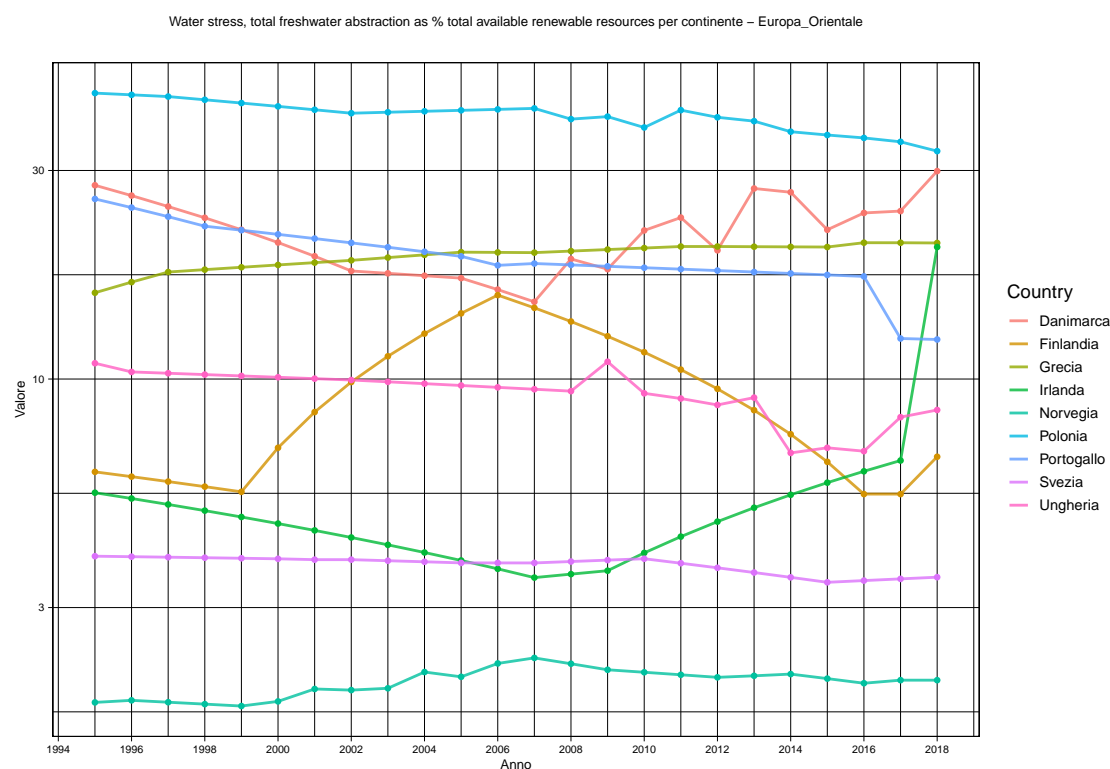


Figura 3.31: Serie storica per i paesi dell'Europa Orientale: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

Nei continenti dell'**Oceania e Africa** (figura 3.32), si riscontrano contrasti marcati. Mentre l'Oceania ha generalmente un basso stress idrico, in Africa il quadro è molto diverso. Il Sud Africa affronta livelli critici di stress idrico, spesso dovuti a scarsità di risorse, infrastrutture inadeguate e gestione inefficace delle risorse idriche.

In generale, questi grafici a serie storica forniscono una panoramica dell'andamento dei livelli di stress idrico nei diversi continenti nel periodo considerato. Sono evidenziate situazioni di aumento, diminuzione o stabilità dello stress idrico, offrendo spunti importanti per l'analisi della gestione delle risorse idriche a livello regionale e globale.

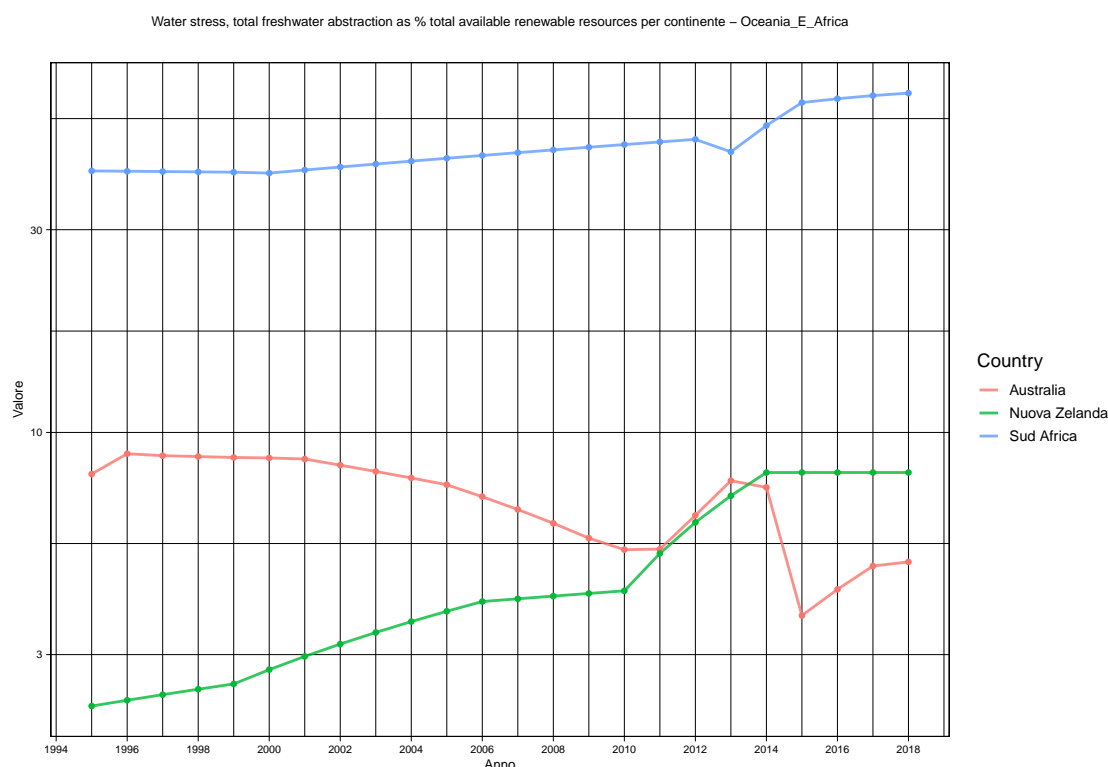


Figura 3.32: Serie storica per i continenti: Oceania e Africa: Stress idrico, prelievo totale di acqua dolce come % delle risorse rinnovabili totali disponibili

3.2.5 Fornitura di energia rinnovabile, % del totale fornitura energetica

Il tasso di fornitura di energia rinnovabile è calcolato come percentuale delle fonti rinnovabili del totale fornitura energetica (Totale Fornitura Energetica (TES)). Le fonti rinnovabili includono idroelettrico, geotermico, solare (termico e fotovoltaico), eolica e energia mareale/ondosa/oceanica, nonché combustibili rinnovabili (biomasse solide, biomasse liquide, biogas) e rifiuti (rifiuti urbani rinnovabili). Il TES (Totale Fornitura Energetica) è calcolato come la somma della produzione nazionale di energia, delle importazioni, e delle variazioni di scorte, al netto delle esportazioni, dei bunker navali internazionali e dei bunker aerei internazionali.

I dati sottostanti sulla "fornitura di energia da fonti rinnovabili e rifiuti (ktoe)" sono ottenuti dal set di dati World - Renewable and Waste Energy Statistics del database International Energy Agency (IEA) Renewables Information Statistics Database. I dati sulla fornitura totale di energia (TES) sono ottenuti dal database IEA sulle statistiche e bilanci energetici mondiali.

Il **boxplot** nella figura 3.33 offre una rappresentazione visiva della distribuzione della variabile "Fornitura di energia rinnovabile come % del totale fornitura energetica" durante il periodo considerato. Questo grafico fornisce importanti informazioni statistiche sulla distribuzione di questa variabile. La **mediana** della distribuzione è di 7.81%, il che indica che la metà dei paesi ha una fornitura di energia rinnovabile inferiore a questo valore. Questa mediana rappresenta un punto di riferimento importante per valutare la situazione dell'uso di energia rinnovabile nei paesi. La **media** è di 13.5%, leggermente superiore alla mediana. Questo suggerisce che ci sono alcuni paesi con valori significativamente superiori alla media, che contribuiscono a tirare verso l'alto

Fornitura di energia rinnovabile, % del totale fornitura energetica	
Media	13.5
Mediana	7.81
Simmetria	1.141139
Curtosi	0.1143298
Frequenza assoluta	744
Massimo	51.5
Minimo	0.3
Deviazione standard	12.59756
Coefficiente di variazione	0.9323549
Varianza	158.6984
Q1	4.60
Q3	20.6

Tabella 3.5: Statistiche descrittive per la variabile: Fornitura di energia rinnovabile, % del totale fornitura energetica

la distribuzione complessiva. La distribuzione è quindi leggermente **asimmetrica positivamente**.

Il boxplot mostra chiaramente la presenza di alcuni **valori estremi superiori**, evidenziati come punti al di fuori dei "baffi" del grafico. Il valore massimo registrato è del 51.5%, il che indica che alcuni paesi dipendono in larga misura da fonti di energia rinnovabile. Questi valori estremi superiori contribuiscono alla presenza di asimmetria positiva nella distribuzione.

La **deviazione standard** è di 12.59756, suggerendo una certa variabilità nei livelli di fornitura di energia rinnovabile tra i paesi considerati. Il **coefficiente di variazione** è 0.9323549, che rappresenta la variabilità relativa rispetto alla media, indica che la variabilità nei livelli di fornitura di energia rinnovabile è significativa rispetto alla media. La **varianza** è di 158.6984, confermando la variabilità nei dati.

I **quartili** (Q1 e Q3) sono rispettivamente 4.60% e 20.6%, indicando la distribuzione dei dati e suggerendo che molti paesi hanno una fornitura di energia rinnovabile relativamente bassa, mentre pochi paesi registrano valori molto alti.

La simmetria positiva della distribuzione, indicata da un valore di **skewness** pari a 1.141139, conferma la presenza di una coda lunga verso valori più alti. Ciò significa che alcuni paesi hanno una fornitura di energia rinnovabile significativamente superiore alla media, influenzando in modo significativo la distribuzione complessiva. La **curtosi** vicina a zero suggerisce che la distribuzione non è né troppo appuntita né troppo piatta rispetto a una distribuzione normale. Questo implica una distribuzione relativamente equilibrata dei dati, nonostante la presenza di alcuni valori estremi.

L'**istogramma** rappresenta la distribuzione della variabile durante il periodo considerato. Questo tipo di grafico è utile per visualizzare la frequenza con cui i diversi valori della variabile si verificano. Sull'asse delle ascisse sono riportati i possibili valori assunti dalla variabile, ossia le percentuali di fornitura di energia rinnovabile. Questo intervallo va dal **minimo valore** dello 0.3% al **massimo valore** del 51.5%. Sull'asse delle ordinate è rappresentata la frequenza con cui ciascun valore si verifica. In altre parole, l'altezza delle barre sull'istogramma rappresenta quanti paesi presentano una determinata percentuale di fornitura di energia rinnovabile.

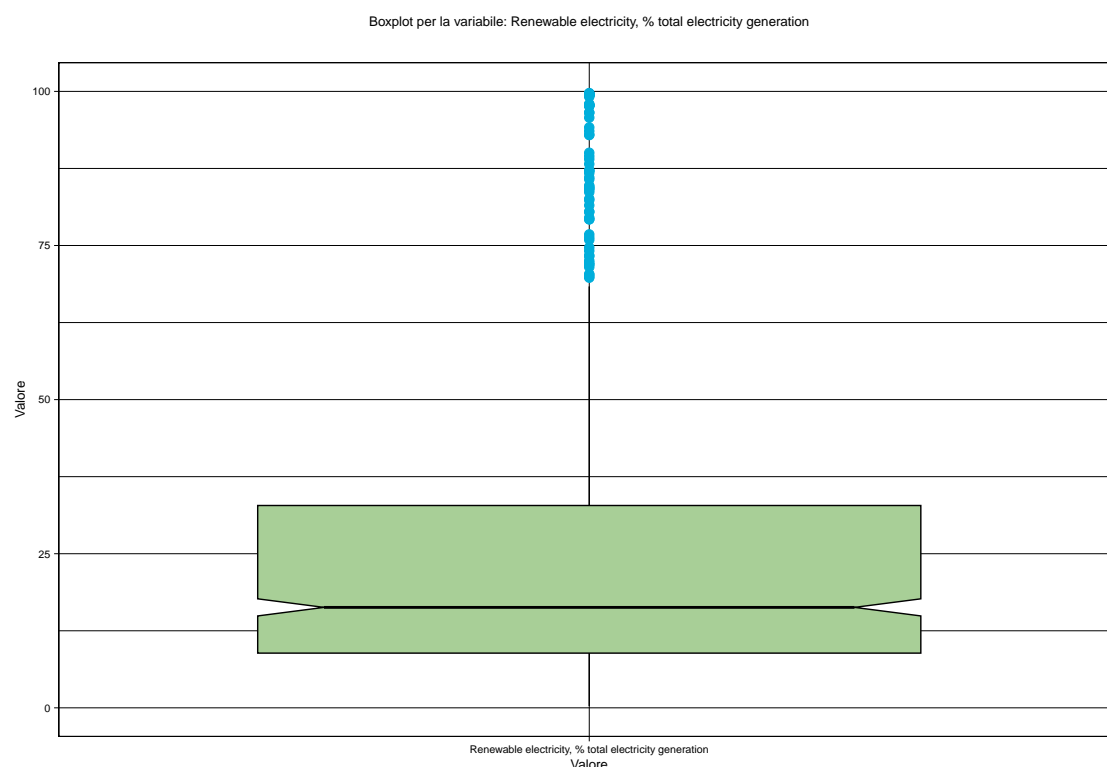


Figura 3.33: Boxplot per la variabile: Fornitura di energia rinnovabile, % del totale fornitura energetica

La maggior parte dei paesi sembra avere una percentuale relativamente bassa di fornitura di energia rinnovabile, concentrata nella parte sinistra dell'istogramma. Questo è evidenziato dalla presenza di una barra molto alta in quella regione. Tuttavia, l'istogramma mostra anche la presenza di alcuni paesi con una fornitura di energia rinnovabile significativamente più elevata, come indicato da una coda lunga verso destra con alcune barre che si estendono in quella direzione. L'istogramma mostra una distribuzione leggermente asimmetrica positiva, in quanto la coda verso destra (valori più alti) è più lunga rispetto alla coda verso sinistra (valori più bassi). Questa asimmetria è coerente con quanto osservato nel boxplot, dove sono presenti valori estremi superiori.

La **funzione di densità** rappresentata nel grafico 3.35 offre una visualizzazione della distribuzione della variabile durante il periodo considerato. Questo tipo di grafico è noto come grafico della densità di probabilità o **kernel density plot**.

Nel grafico:

- Sull'asse delle ascisse (orizzontale) sono riportati i possibili valori assunti dalla variabile, ovvero le percentuali di fornitura di energia rinnovabile. Questo intervallo va dal **minimo valore** dello 0.3% al **massimo valore** del 51.5%.
- Sull'asse delle ordinate (verticale) è rappresentata la **densità di probabilità**, che indica la probabilità relativa di trovare un'osservazione in una determinata fascia di valori.

La curva nel grafico mostra la forma della distribuzione della variabile. In questo caso, la curva è **asimmetrica positiva**, con una coda lunga che si estende verso valori più alti. Questo con-

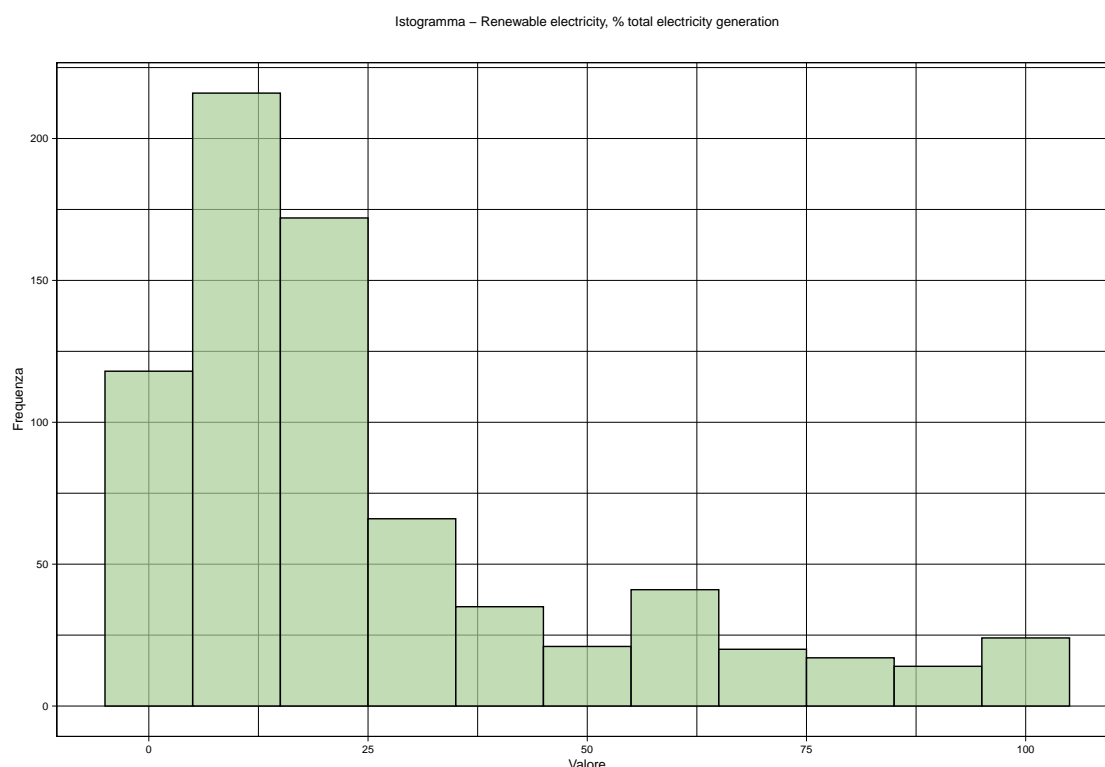


Figura 3.34: Istogramma per la variabile: Fornitura di energia rinnovabile, % del totale fornitura energetica

ferma quanto osservato in precedenza sulla simmetria positiva dei dati. Le regioni più alte della curva corrispondono alle percentuali di fornitura di energia rinnovabile in cui si trovano più paesi. Quindi, le percentuali intorno alla **mediana** sono le più comuni. La curva è più alta nella parte sinistra, indicando che molte osservazioni si concentrano nelle percentuali più basse di fornitura di energia rinnovabile. Questo è coerente con quanto visto nei quartili Q1 e Q3 della distribuzione. La densità diminuisce man mano che ci spostiamo verso le percentuali più alte, ma la coda lunga indica che ci sono comunque paesi con percentuali molto elevate di energia rinnovabile.

La figura 3.36 illustra chiaramente le differenze nei livelli e nelle tendenze della produzione di energia rinnovabile tra i paesi del **continente americano**. Mentre il **Brasile** mostra un impegno crescente verso l'energia sostenibile, il **Canada** dimostra stabilità, e gli altri paesi, pur facendo progressi, mantengono ancora una quota significativa di energia da fonti non rinnovabili. Questi risultati evidenziano le diverse strategie energetiche adottate dai paesi americani e sottolineano l'importanza di continuare a promuovere lo sviluppo delle energie rinnovabili.

Nella figura 3.37 è rappresentata la serie storica della produzione di energia rinnovabile per i paesi del **continente Asiatico** nel periodo considerato. L'Indonesia, l'India e la Cina mostrano una tendenza decrescente nella quota di energia rinnovabile. Questo potrebbe essere attribuito a diversi fattori, tra cui un aumento dell'approvvigionamento totale di energia a causa dell'industrializzazione in corso in questi paesi. In molti casi, un rapido sviluppo industriale può portare a un aumento della domanda di energia, che potrebbe essere soddisfatta principalmente da fonti non rinnovabili come il carbone. Questo può influire sulla quota di energia rinnovabile. Il **Giappone** mostra una

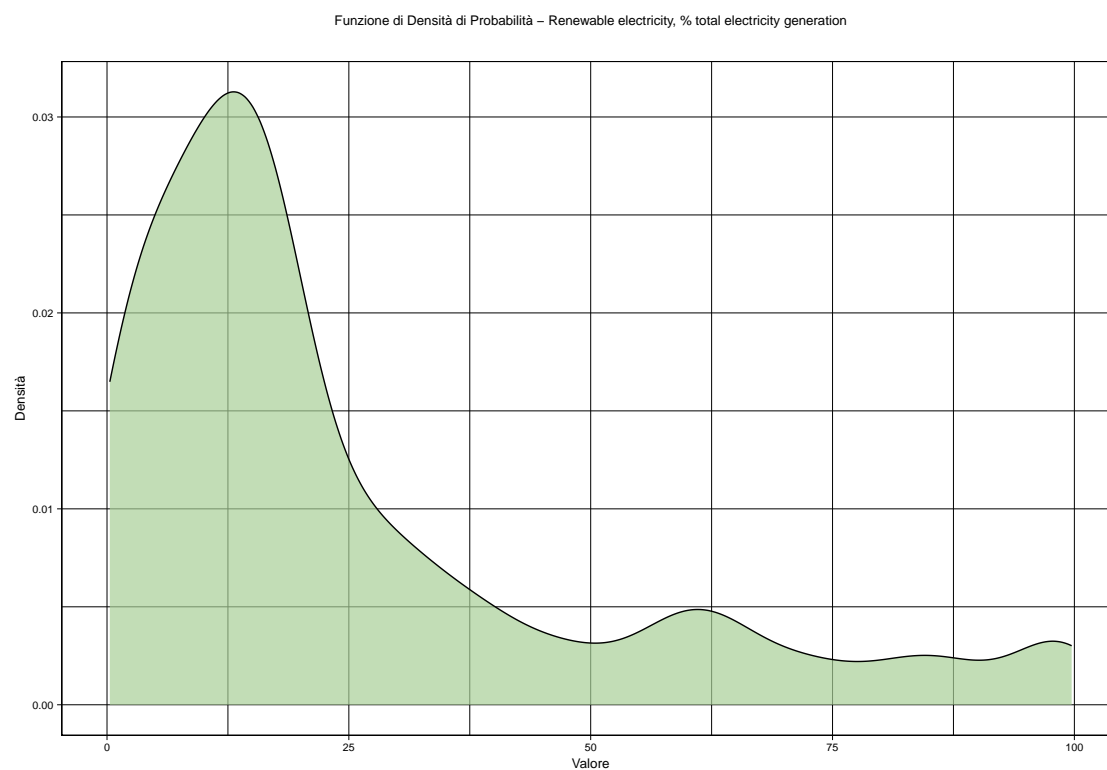


Figura 3.35: Funzione di densità per la variabile: Fornitura di energia rinnovabile, % del totale fornitura energetica

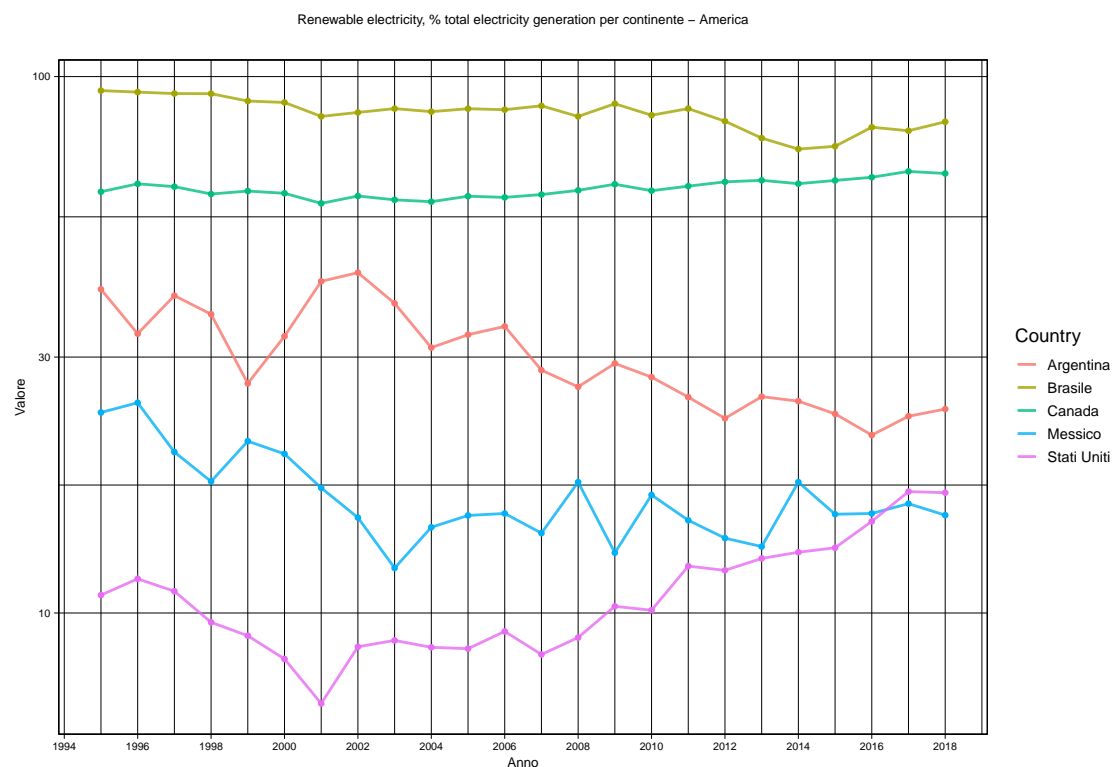


Figura 3.36: Serie storica per il continente Americano: Fornitura di energia rinnovabile, % del totale fornitura energetica

leggera crescita pur mantenendo la quota attorno al 5%, mentre Russia e Corea del Sud mantengono una percentuale costante prossima allo zero.

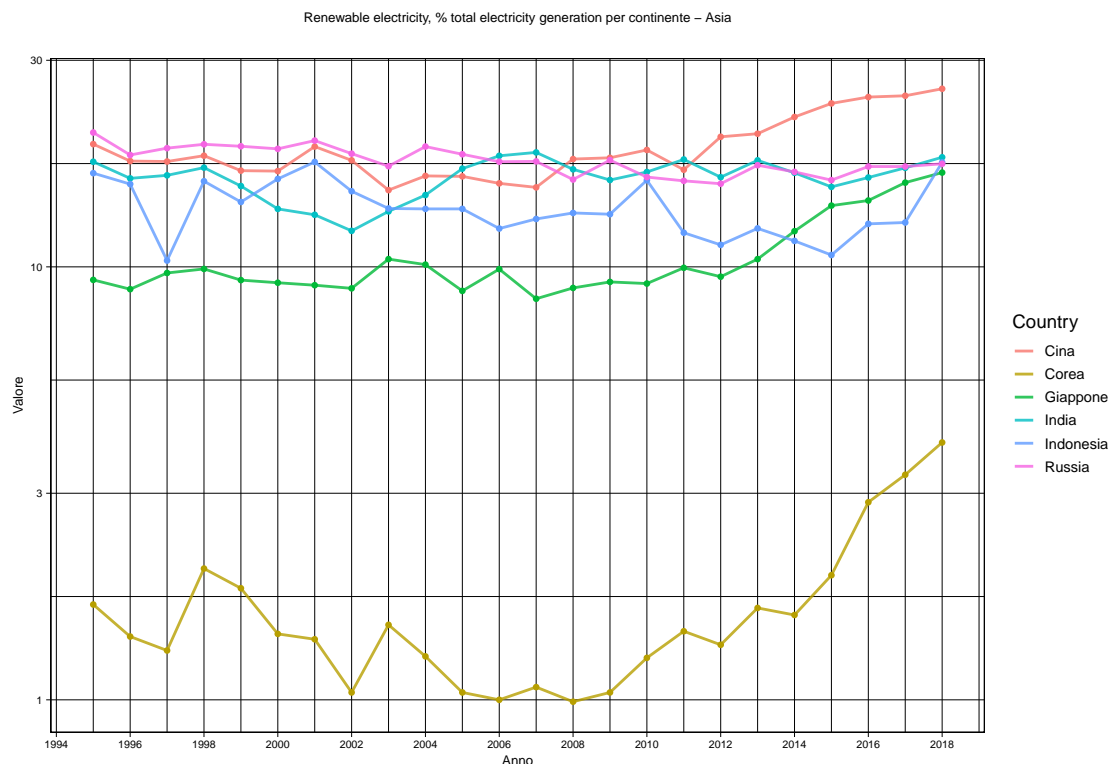


Figura 3.37: Serie storica per il continente Asiatico: Fornitura di energia rinnovabile, % del totale fornitura energetica

La figura 3.38 mostra la serie storica della produzione di energia rinnovabile per i paesi dell'Europa Occidentale durante il periodo considerato.

Tutti i paesi considerati mostrano un aumento della quota di energia sostenibile. L'Italia parte da una quota energetica del 5% per arrivare al 15%, mentre la Germania parte da una quota vicina all'1% arrivando a poco più dell'11%. La Francia e la Spagna mostrano una crescita meno costante rispetto all'Italia, con periodi di decrescita. I restanti paesi partono da una quota energetica vicina allo 0% per arrivare a quote variegiate tra il 5% e il 10%. In generale, si evidenzia l'impegno di queste nazioni verso l'energia sostenibile.

La figura 3.39 presenta la serie storica della produzione di energia rinnovabile per paesi dell'Europa Orientale nel periodo dal 1995 al 2018.

Anche in questo caso vi è una crescita per la quota dell'energia sostenibile per i paesi considerati. In particolare, paesi come Irlanda, Ungheria, Polonia e Grecia mostrano una crescita limitata, dal circa 0% per arrivare a circa il 10%. La Danimarca ha effettuato un considerevole aumento della quota di energia sostenibile, passando dal circa 10% fino al 30%. La Finlandia, il Portogallo e la Svezia mostrano una crescita della quota di energia sostenibile più piccola rispetto alla Danimarca e soprattutto meno costante, con periodi di decrescita. La Norvegia mostra una quota considerevole di energia sostenibile di circa il 50%, ma non mostra incrementi significativi; al contrario, mostra un andamento altalenante anche se alla fine della serie storica conferma la sua

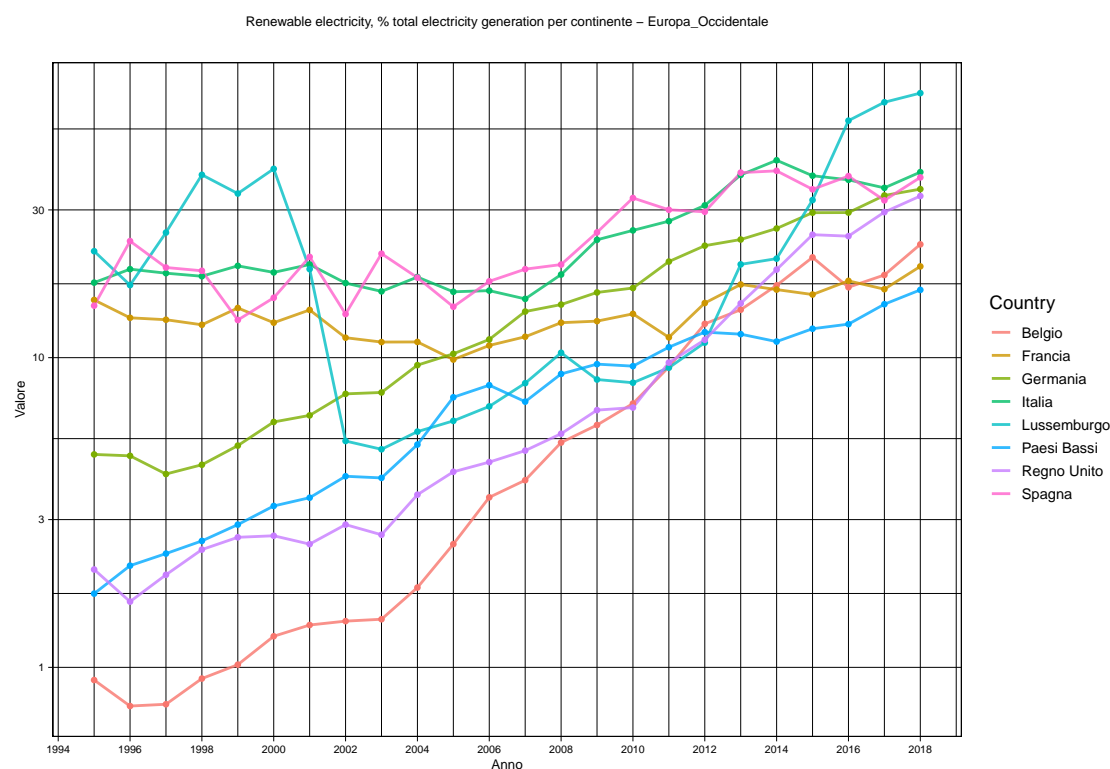


Figura 3.38: Serie storica per i paesi dell'Europa Occidentale: Fornitura di energia rinnovabile, % del totale fornitura energetica

quota energetica. In generale, per i paesi dell'Europa Orientale si osserva una quota di energia rinnovabile maggiore rispetto all'Europa Occidentale. In generale, questi dati riflettono l'impegno crescente verso l'energia sostenibile in Europa, con alcuni paesi che stanno compiendo progressi significativi, mentre altri affrontano sfide nella transizione verso fonti energetiche più pulite.

La figura 3.40 presenta la serie storica della produzione di energia rinnovabile per i paesi dell'Oceania e dell'Africa, nel periodo dal 1995 al 2018.

Per la **Nuova Zelanda** possiamo osservare, nonostante alcuni periodi di decrescita, una crescita della percentuale di energia rinnovabile. Al contrario, il **Sud Africa** mostra una decrescita dal circa 10% a circa lo 0%. Ovviamente, la decrescita osservata può anche essere causata da un aumento dell'approvvigionamento totale di energia legata a una maggiore industrializzazione del paese.

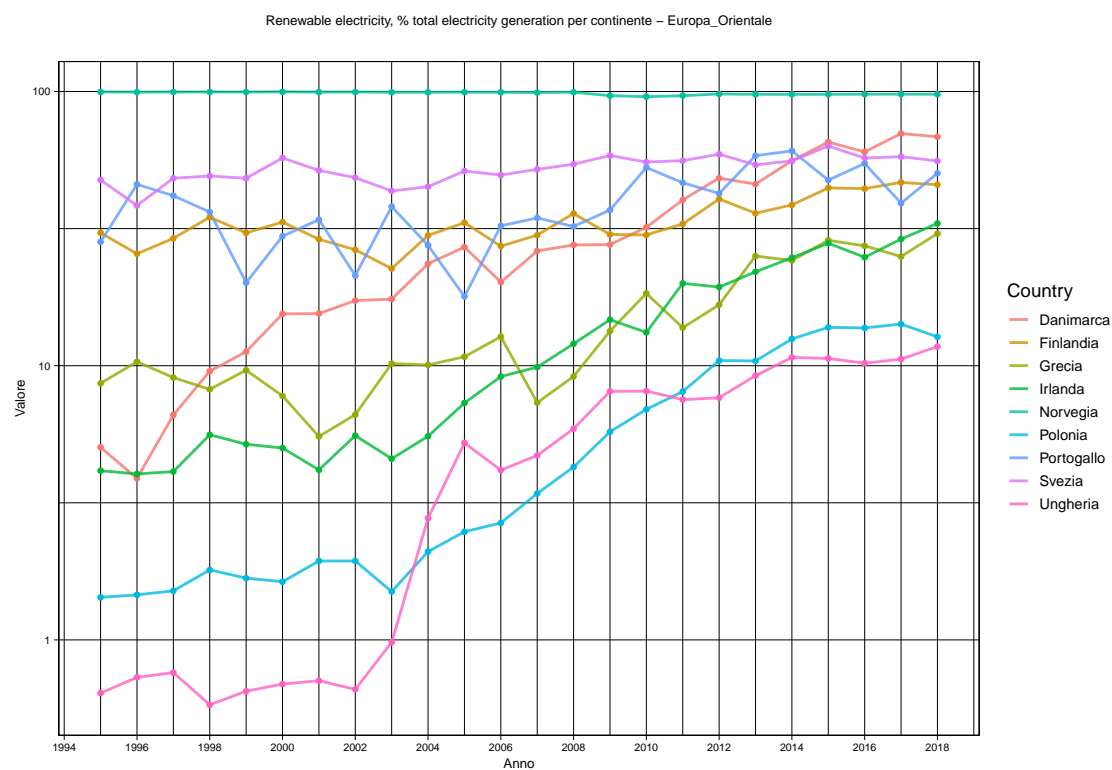


Figura 3.39: Serie storica per i paesi dell'Europa Orientale: Fornitura di energia rinnovabile, % del totale fornitura energetica

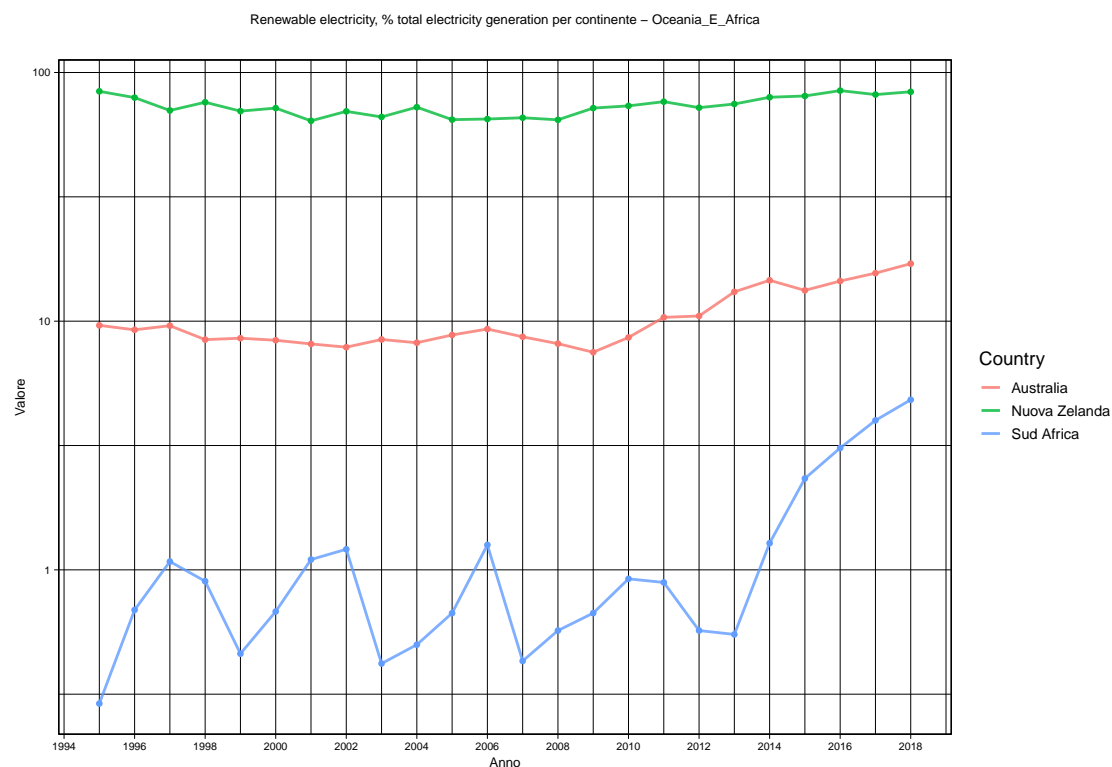


Figura 3.40: Serie storica per i paesi dell'Oceania e Africa: Fornitura di energia rinnovabile, % del totale fornitura energetica

3.2.6 Elettricità Rinnovabile, % della Totale Produzione Elettrica

L'energia elettrica rinnovabile è calcolata come una quota di energia rinnovabile nella produzione di energia elettrica (%). Le energie rinnovabili includono: idroelettrico, geotermico, solare (termico e fotovoltaico), eolico ed energia mareo-onda-oceanica, nonché combustibili rinnovabili (biomasse solide, biomasse liquide, biogas) e rifiuti (rifiuti urbani rinnovabili).

La quota di energia elettrica rinnovabile nella produzione totale di energia elettrica è un indicatore importante dell'utilizzo di fonti di energia pulite e sostenibili. La crescita della quota di energia elettrica rinnovabile è un obiettivo chiave della politica energetica globale, in quanto contribuisce a ridurre le emissioni di gas serra e a migliorare la sicurezza energetica.

La crescita della quota di energia elettrica rinnovabile è guidata da una serie di fattori, tra cui:

1. L'aumento dei costi dei combustibili fossili: La crescente domanda di energia e la scarsità di combustibili fossili stanno facendo aumentare i costi di questi combustibili. Questo sta rendendo le fonti di energia rinnovabili più competitive, in quanto i loro costi sono diminuiti negli ultimi anni.
2. L'aumento della consapevolezza ambientale: La crescente preoccupazione per il cambiamento climatico e altri problemi ambientali sta spingendo i governi e le imprese a investire in fonti di energia pulite e sostenibili.
3. Lo sviluppo di nuove tecnologie: Le nuove tecnologie, come i pannelli solari più efficienti e le turbine eoliche più potenti, stanno rendendo le fonti di energia rinnovabili più convenienti e accessibili.

La crescita della quota di energia elettrica rinnovabile è un passo importante verso un futuro energetico più pulito e sostenibile.

Elettricità Rinnovabile, % della Totale Produzione Elettrica	
Media	25.4
Mediana	16.3
Simmetria	1.465705
Curtosi	1.251648
Frequenza assoluta	744
Massimo	99.7
Minimo	0.29
Deviazione standard	25.31156
Coefficiente di variazione	0.9965889
Varianza	640.6752
Q1	8.88
Q3	32.8

Tabella 3.6: Statistiche descrittive per la variabile: Elettricità Rinnovabile, % della Totale Produzione Elettrica

Il **boxplot** rappresenta la distribuzione della variabile "Elettricità Rinnovabile come % della Totale Produzione Elettrica" durante il periodo considerato. Questo tipo di grafico è utile per visualizzare la distribuzione dei dati e individuare eventuali valori estremi. La linea orizzontale all'interno del box rappresenta la mediana, che è pari al 16.3%. Questo indica che la metà dei paesi ha una

percentuale di produzione di elettricità rinnovabile inferiore a questo valore.

La **deviazione standard** è di 25.31156, il che indica una certa variabilità nei livelli di produzione di elettricità rinnovabile tra i paesi. La **varianza** è di 640.6752, confermando la variabilità nei dati. Il **coefficiente di variazione** (0.9965889) rappresenta la variabilità relativa rispetto alla media, indicando che la variabilità nei livelli di produzione di elettricità rinnovabile è significativa rispetto alla media.

Il **boxplot** mostra Intervallo Interquartile (IQR), che va dal Primo Quartile (Q1) al Terzo Quartile (Q3). In questo caso, Q1 è pari all'8.88% e Q3 è pari al 32.8%. L'Intervallo Interquartile (IQR) rappresenta la zona centrale dei dati, dove si concentra la maggior parte delle osservazioni. Le linee verticali (baffi) si estendono dall'Intervallo Interquartile (IQR) fino ai valori minimi e massimi che rientrano entro 1,5 volte l'Intervallo Interquartile (IQR). Ogni punto al di fuori di questi baffi è considerato un valore anomalo o un valore estremo. Nel grafico, ci sono diversi valori anomali superiori, indicando che alcuni paesi hanno una percentuale molto alta di produzione di elettricità rinnovabile.

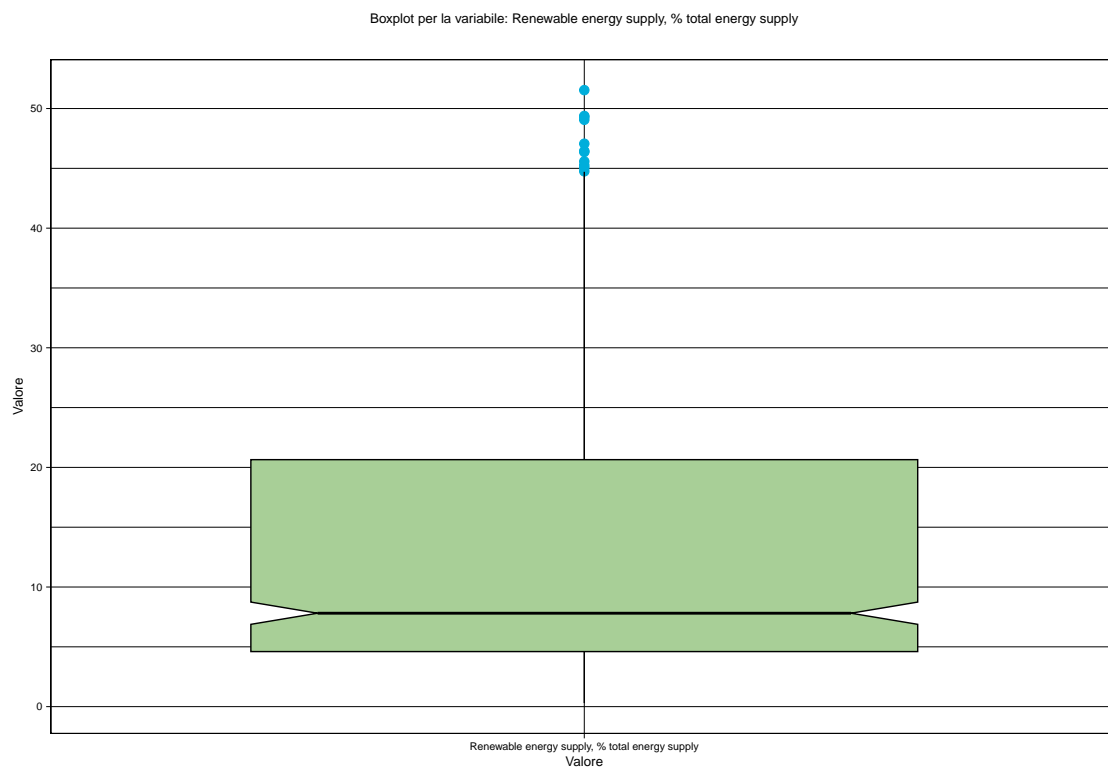


Figura 3.41: Boxplot per la variabile: Elettricità Rinnovabile, % della Totale Produzione Elettrica

La **simmetria positiva della distribuzione**, indicata da **skewness** = 1.465705, suggerisce che la coda lunga si estende verso valori più alti, confermando la presenza di quei paesi con una produzione di elettricità rinnovabile significativamente superiore alla media. La **curtosi**, indicata da **kurtosis** = 1.251648, suggerisce una certa "pesantezza" delle code della distribuzione, il che significa che ci sono più valori estremi rispetto a una distribuzione normale.

L'**istogramma** 3.42 rappresenta la distribuzione della variabile "Elettricità Rinnovabile come % della Totale Produzione Elettrica" durante il periodo considerato. Questo tipo di grafico è utile per visualizzare la frequenza con cui si verificano determinati intervalli di valori nella distribuzione dei dati. Sull'asse delle ascisse sono rappresentati gli intervalli di valori, mentre sull'asse delle ordinate è rappresentata la frequenza delle osservazioni che rientrano in ciascun intervallo.

La **distribuzione è positivamente asimmetrica**, il che significa che ci sono più paesi con una percentuale relativamente bassa di produzione di elettricità rinnovabile rispetto a quelli con una percentuale molto alta. La coda lunga verso destra indica la presenza di alcuni paesi con una produzione di elettricità rinnovabile molto alta, contribuendo all'asimmetria positiva. La classe con la frequenza più alta sembra essere nell'intervallo compreso tra il 0% e il 10%, il che indica che molti paesi hanno una produzione di elettricità rinnovabile inferiore al 10% della loro produzione totale di energia elettrica. Ci sono alcune classi più ampie, come quella tra il 20% e il 30%, che potrebbe indicare un gruppo di paesi con una produzione moderata di elettricità rinnovabile. Gli intervalli di valori più elevati, come quelli superiori al 40% o 50%, sono meno frequenti, ma mostrano la presenza di paesi con una produzione di elettricità rinnovabile molto significativa rispetto alla loro produzione totale.

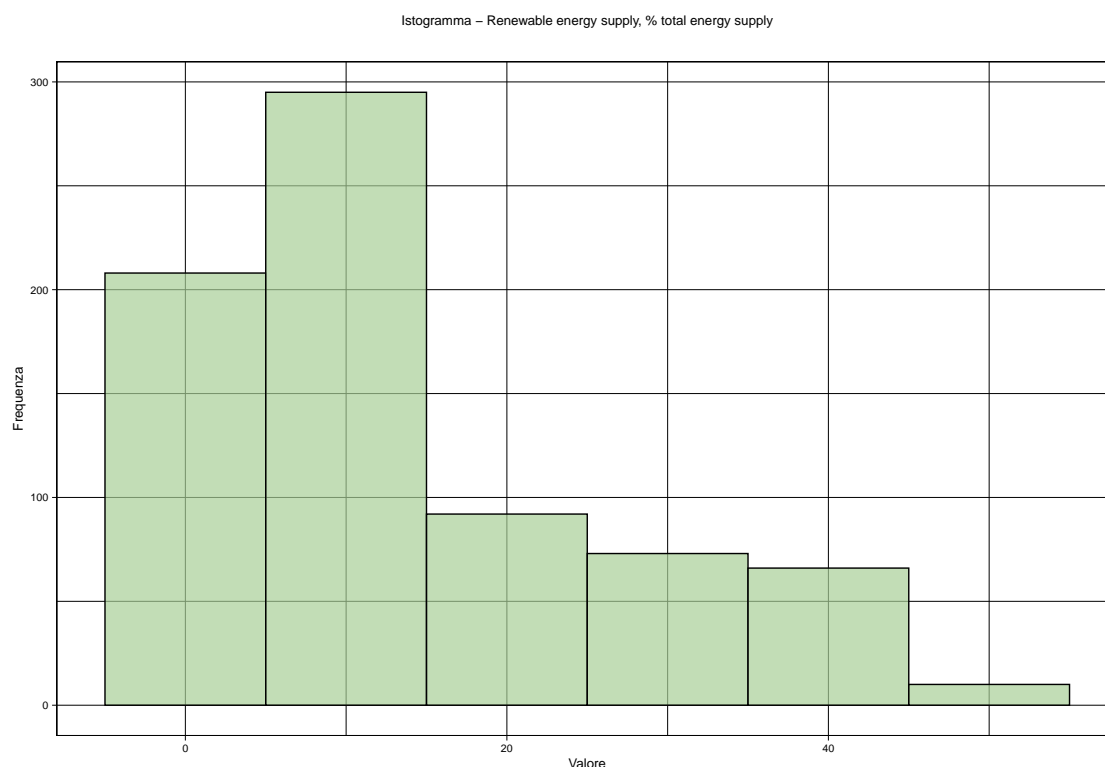


Figura 3.42: Istogramma per la variabile: Elettricità Rinnovabile, % della Totale Produzione Elettrica

La funzione di densità 3.43 di probabilità mostra la forma della distribuzione, confermando nuovamente l'asimmetria positiva con una coda lunga verso valori più alti.

Nelle figure 3.44, 3.45, 3.46, 3.47 e 3.48 è possibile osservare la serie storica della variabile per tutti

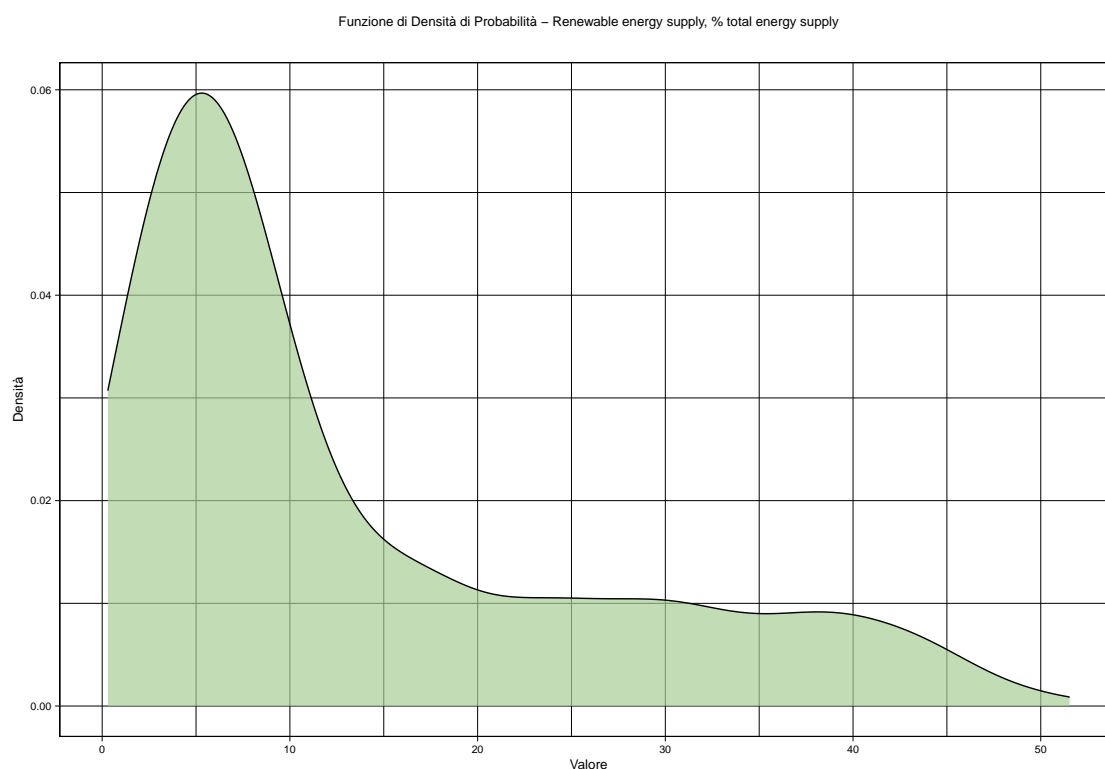


Figura 3.43: Funzione di densità per la variabile: Elettricità Rinnovabile, % della Totale Produzione Elettrica

i paesi considerati, nel periodo 1995/2018. Sull'asse delle ascisse sono presenti gli anni mentre sull'asse delle ordinate sono presenti i valori assunti dalla variabile. Ciascuna linea all'interno del grafico rappresenta la serie storica di uno specifico paese.

L'analisi della figura 3.44, che mostra le tendenze storiche nell'utilizzo di energia elettrica nel **continente americano**, rivela una situazione estremamente diversificata tra i diversi paesi, ognuno con le proprie caratteristiche e sfide uniche.

Il **Messico e gli Stati Uniti** mostrano una quota relativamente bassa di energia elettrica proveniente da fonti rinnovabili. Questo dato potrebbe essere attribuito a vari fattori. Entrambi i paesi hanno economie altamente sviluppate con una grande dipendenza da fonti energetiche tradizionali come petrolio, gas naturale e carbone. Il **Canada**, d'altra parte, si distingue per una quota costantemente alta di energia elettrica proveniente da fonti rinnovabili, mantenendosi intorno al 60%. Questo è probabilmente dovuto alle abbondanti risorse naturali del paese, come l'energia idroelettrica, che hanno permesso al Canada di sfruttare efficacemente queste fonti per la produzione di energia elettrica. Il **Brasile** presenta un caso interessante con una leggera decrescita nell'utilizzo di energia rinnovabile, passando da circa il 90% al 77%. Questa tendenza potrebbe riflettere diversi fattori. Il Brasile ha una lunga storia nell'uso di energie rinnovabili, in particolare l'idroelettrico e i biocarburanti. Tuttavia, la decrescita potrebbe indicare una dipendenza crescente da fonti energetiche non rinnovabili, forse dovuta a un aumento della domanda energetica complessiva.

L'analisi della figura 3.45, che presenta le serie storiche dell'utilizzo di energia elettrica in alcuni

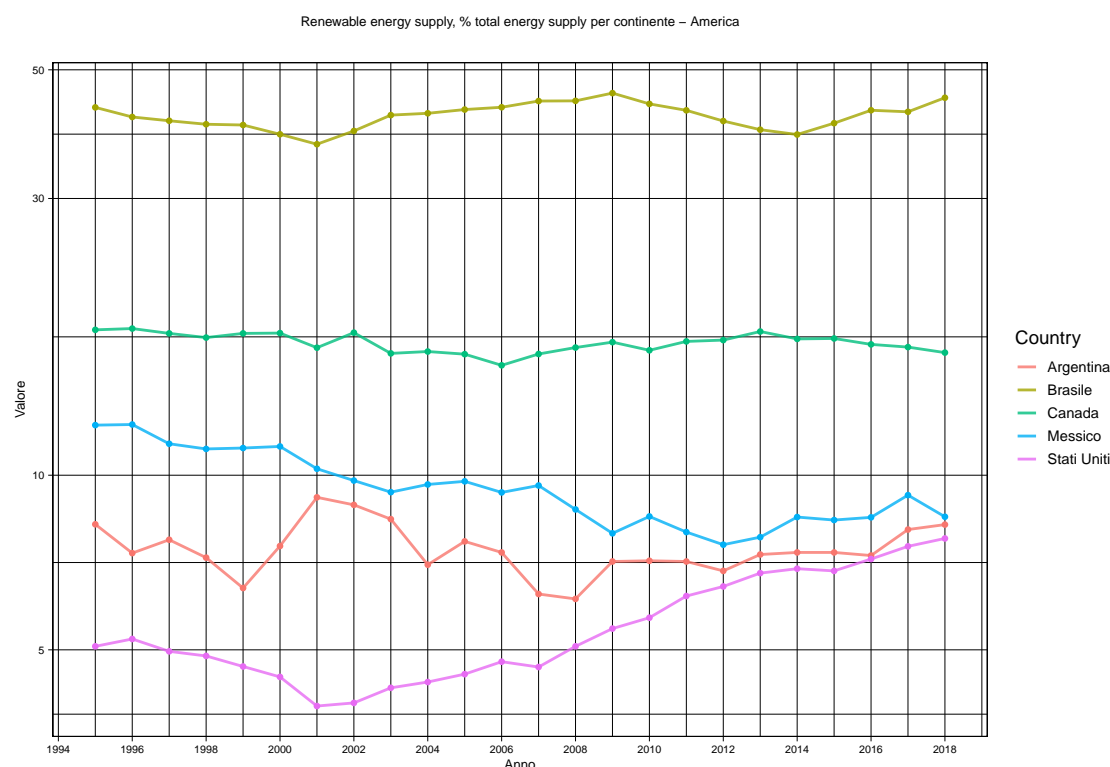


Figura 3.44: Serie storica per il continente Americano: Elettricità Rinnovabile, % della Totale Produzione Elettrica

paesi asiatici, rivela una serie di tendenze interessanti e eterogenee. La Corea mostra una relativa stabilità nella sua quota di energia elettrica, oscillando tra lo 0% e il 5%. Questo potrebbe indicare un approccio cauto o graduale verso l'integrazione di energie rinnovabili nel suo mix energetico. La stabilità potrebbe anche suggerire una forte dipendenza da fonti energetiche tradizionali.

La **Cina** mostra un cambiamento significativo e un rapido aumento nell'utilizzo di energia rinnovabile a partire dal 2010, contrapponendosi a un andamento precedente più discontinuo. Questo notevole cambiamento potrebbe riflettere un impegno politico e finanziario crescente verso la sostenibilità energetica, specialmente considerando le crescenti preoccupazioni ambientali e le pressioni internazionali. **Russia, India e Indonesia** mostrano un andamento molto discontinuo caratterizzato da fluttuazioni tra periodi di crescita e decrescita. Questo potrebbe riflettere l'instabilità delle politiche energetiche, le fluttuazioni economiche o le sfide tecniche nella gestione e nello sviluppo di infrastrutture per l'energia rinnovabile. In particolare, in paesi con rapida crescita economica come l'India e l'Indonesia, la crescente domanda di energia potrebbe portare a una maggiore varietà nelle fonti energetiche utilizzate.

Il **Giappone**, con un aumento più regolare dall'10% al 15%, mostra un impegno costante e progressivo verso l'energia rinnovabile. Questo approccio graduale potrebbe essere il risultato di politiche energetiche ben pianificate e di un impegno per bilanciare la necessità di energia affidabile con quella della sostenibilità ambientale. Il Giappone, in particolare dopo il disastro di Fukushima, ha rivisto le sue politiche energetiche, spostando l'attenzione dalle fonti nucleari a quelle rinnovabili.

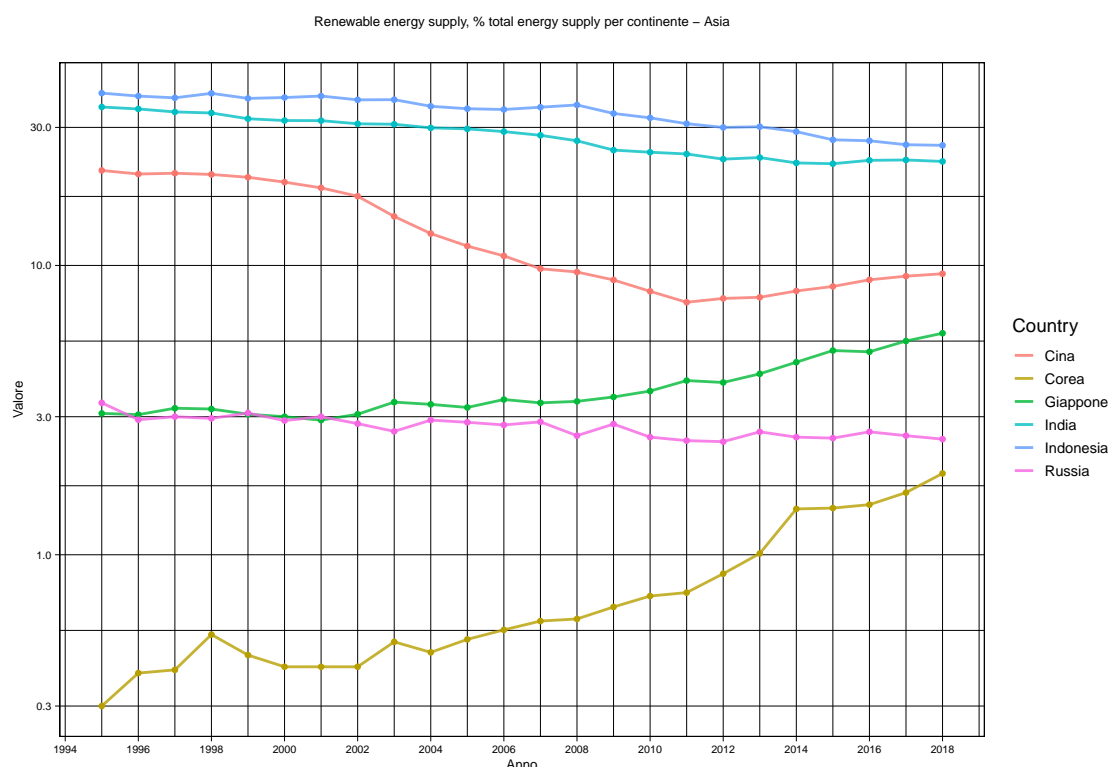


Figura 3.45: Serie storica per il continente Asiatico: Elettricità Rinnovabile, % della Totale Produzione Elettrica

La figura 3.46 offre un'analisi dettagliata delle tendenze nell'utilizzo di energia rinnovabile nei paesi dell'Europa Occidentale. Diversi paesi mostrano pattern distinti, riflettendo le varie strategie e sfide nella transizione verso fonti di energia più sostenibili.

Belgio, Regno Unito, Paesi Bassi e Germania hanno mostrato un percorso simile, partendo da basse percentuali di energia rinnovabile e crescendo in modo abbastanza costante. I valori finali di 22% per il Belgio, 30% per il Regno Unito, 18% per i Paesi Bassi e 32% per la Germania indicano un impegno crescente e sostenuto verso l'energia rinnovabile. La crescita costante potrebbe essere il risultato di politiche governative mirate, investimenti in infrastrutture rinnovabili e una crescente consapevolezza ambientale tra la popolazione. La Germania, in particolare, è nota per il suo *Energiewende*, un ambizioso piano di transizione energetica.

La **Francia** mostra una situazione più stabile, con fluttuazioni tra il 15% e il 20%. Questa relativa stabilità potrebbe essere dovuta al forte impegno della Francia nell'energia nucleare, che ha rappresentato una significativa quota del suo mix energetico. Le fluttuazioni osservate potrebbero riflettere vari sforzi nel tempo per integrare più fonti rinnovabili. La **Spagna e l'Italia** presentano andamenti più variabili. La Spagna, oscillando tra il 18% e il 40%, mostra un impegno significativo verso l'energia rinnovabile, sebbene con periodi di crescita e decrescita. Questo potrebbe riflettere variazioni nelle politiche energetiche o nelle condizioni di mercato. L'Italia mostra un percorso paragonabile, suggerendo una tendenza simile nel sud dell'Europa verso l'adozione di energie rinnovabili, nonostante alcune incertezze.

Il **Lussemburgo** è un caso eccezionale, con un'impennata iniziale della quota di energia rinno-

vabile dal 20% al 40%, seguita da una decrescita e poi un'altra crescita significativa oltre il 70%. Questo andamento potrebbe riflettere specifiche iniziative o cambiamenti politici nel paese. La piccola dimensione e l'economia relativamente flessibile del Lussemburgo potrebbero facilitare una transizione più rapida e dinamica verso fonti di energia rinnovabile.

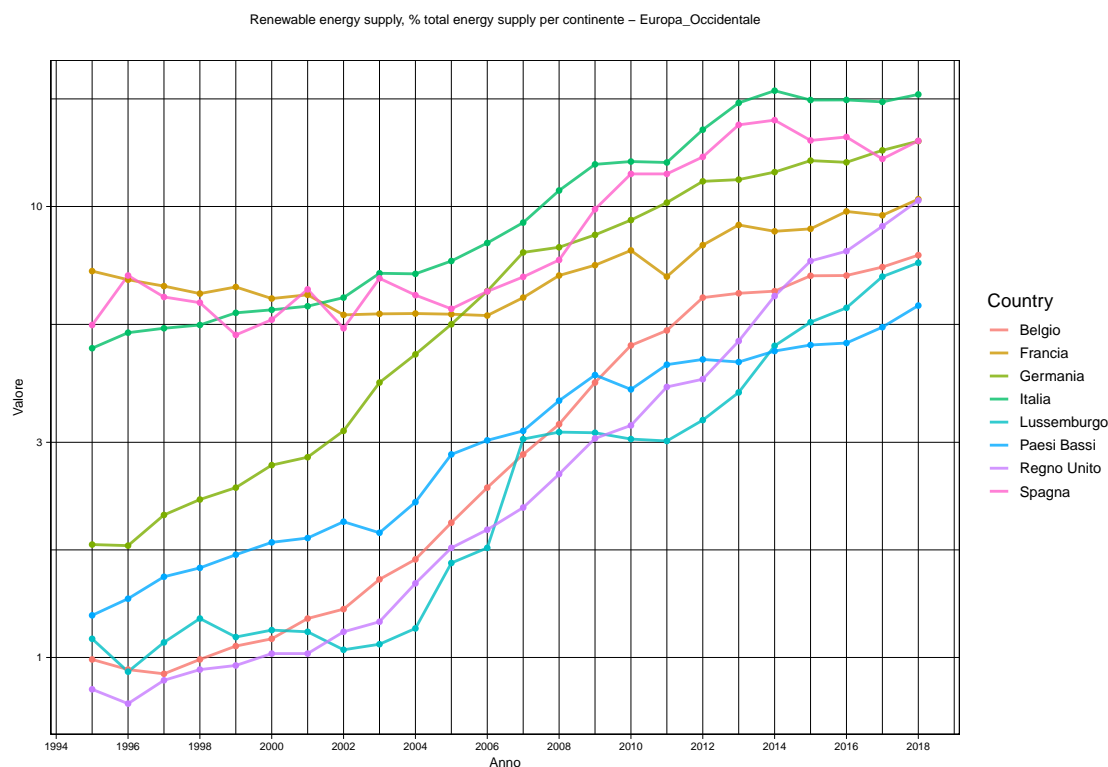


Figura 3.46: Serie storica per i paesi dell'Europa Occidentale: Elettricità Rinnovabile, % della Totale Produzione Elettrica

La figura 3.47 fornisce un quadro dettagliato sulle tendenze dell'utilizzo dell'energia rinnovabile nei paesi dell'Europa Orientale. L'analisi delle diverse traiettorie seguite da questi paesi nell'adozione dell'energia rinnovabile rivela una varietà di approcci e risultati.

La **Norvegia** emerge come un caso eccezionale, mantenendo una quota pari quasi al 100% fino al 2008, con un leggero calo successivo ma un rapido ritorno vicino a questa soglia nel 2018, attestandosi intorno al 98%. Questo andamento sottolinea l'impegno e la capacità della Norvegia di sfruttare le energie rinnovabili, soprattutto l'idroelettrico, che è abbondante nel paese. La lieve fluttuazione osservata potrebbe essere dovuta a variazioni nella produzione idroelettrica o a piccoli aumenti nel consumo di altre fonti energetiche.

La **Danimarca** ha mostrato un incremento notevole, passando dall'1% a circa il 70%. Questo aumento rappresenta un cambiamento significativo nella politica energetica del paese, con un chiaro focus sullo sviluppo sostenibile e sull'energia eolica, per la quale la Danimarca è nota. Questa crescita riflette l'efficacia delle politiche e degli investimenti in infrastrutture rinnovabili.

Sia l'**Irlanda** che la **Grecia** partono da basse percentuali, con l'Irlanda che passa dall'1% al 30%

e la Grecia dal 4% al 20%, sebbene con un andamento più fluttuante. Questi incrementi indicano un impegno crescente verso le energie rinnovabili, ma anche la presenza di sfide e limitazioni nel loro percorso di transizione energetica.

L'**Ungheria e la Polonia** mostrano entrambe un trend di crescita lenta ma costante nell'uso delle energie rinnovabili, con l'Ungheria che raggiunge circa il 12%. Questo può riflettere un approccio graduale e metodico alla transizione energetica, forse influenzato da fattori economici e infrastrutturali.

La **Finlandia e il Portogallo** presentano un andamento altalenante caratterizzato da rapidi periodi di crescita e decrescita. Questa volatilità potrebbe essere attribuita a cambiamenti nelle politiche energetiche, a investimenti variabili nelle energie rinnovabili o a fattori esterni come le condizioni climatiche che influenzano la produzione di energia, soprattutto per fonti come l'eolico e il solare.

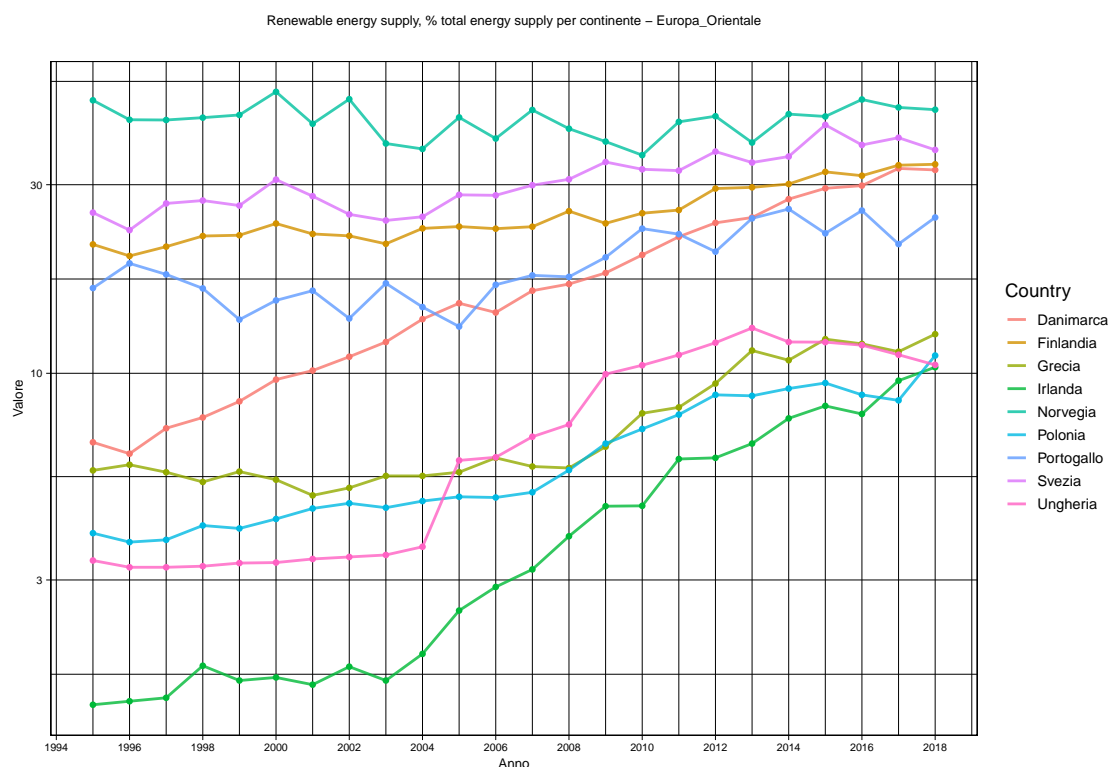


Figura 3.47: Serie storica per i paesi dell'Europa Orientale: Elettricità Rinnovabile, % della Totale Produzione Elettrica

L'analisi delle tendenze nella produzione di energia rinnovabile in Sud Africa, Australia e Nuova Zelanda rivela distinti modelli di sviluppo e adozione di energie sostenibili in questi paesi. Il **Sud Africa** mostra un andamento costante con valori prossimi allo zero per cento nel periodo considerato. Questo indica che la quota di energia rinnovabile rispetto al totale è rimasta molto bassa, suggerendo una forte dipendenza da fonti di energia non rinnovabili. L'**Australia** ha mantenuto un andamento costante intorno al 10% fino al 2012, dopodiché si osserva una crescita fino a circa il 18%. Questo aumento post-2012 potrebbe essere il risultato di un cambiamento nelle politi-

che energetiche, forse stimolato da una crescente consapevolezza ambientale o dalla necessità di rispondere ai cambiamenti climatici. La **Nuova Zelanda** presenta un andamento fluttuante, ma complessivamente mantiene un valore alto, pari all'80%, anche se ci sono stati periodi di crescita e decrescita, con valori che scendono fino al 60%. Questo modello altalenante potrebbe riflettere l'impiego di una varietà di fonti di energia rinnovabile, la cui produzione può variare in base a fattori ambientali e stagionali. La Nuova Zelanda ha un forte impegno verso le energie rinnovabili, specialmente l'idroelettrica e la geotermica, che contribuiscono significativamente al suo mix energetico. Le fluttuazioni osservate potrebbero essere legate alla variabilità delle condizioni ambientali che influenzano la produzione di energia da queste fonti.

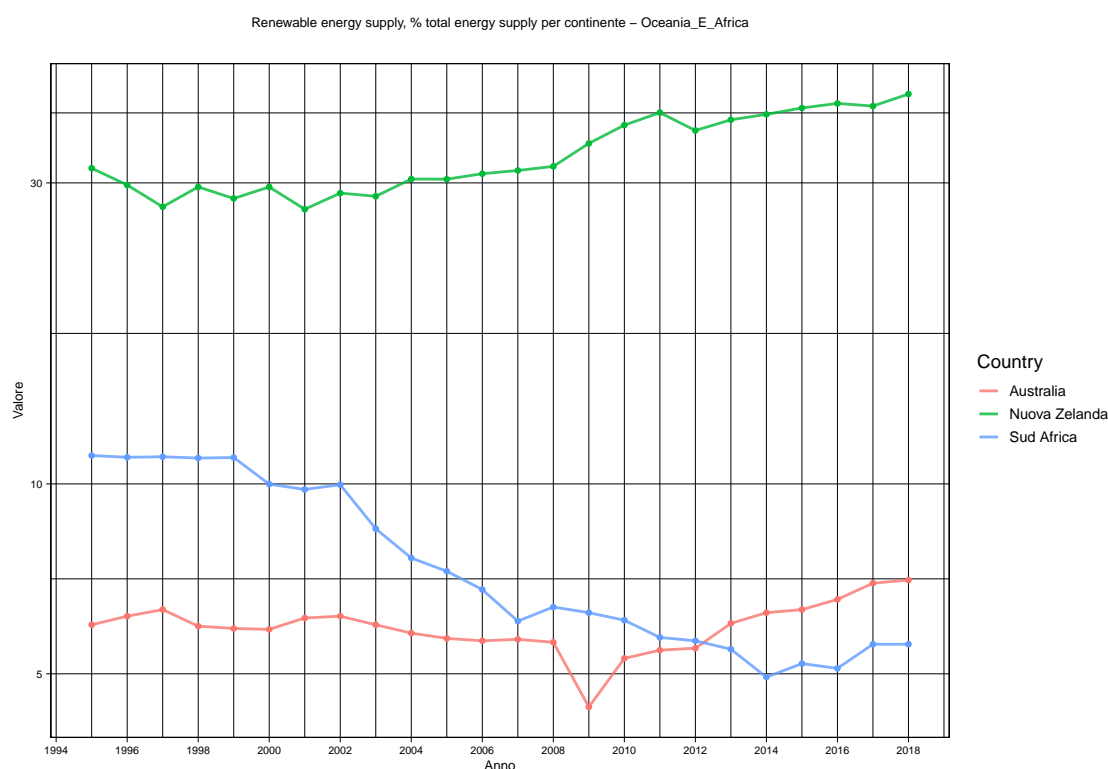


Figura 3.48: Serie storica per i paesi dell'Oceania e dell'Africa: Elettricità Rinnovabile, % della Totale Produzione Elettrica

3.2.7 Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

La mortalità da esposizione a Particulate Matter 2.5 (PM2.5) ambientale è espressa in decessi per milione di abitanti. Secondo (Organizzazione Mondiale della Sanità (OMS)), l'esposizione alle particelle fini sospese nell'aria (Particulate Matter 2.5 (PM2.5)) ha potenzialmente gli effetti avversi più significativi sulla salute rispetto ad altri inquinanti atmosferici. Queste particelle possono essere inalate e causare gravi problemi di salute, tra cui malattie respiratorie e cardiovascolari, con effetti particolarmente gravi su bambini e anziani.

I dati e ulteriori dettagli sulla metodologia possono essere ottenuti dal Global Burden of Disease Study 2019 Results (Studio Globale del Carico di Malattia 2019) e dal dataset sulla mortalità.

Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)	
Media	368
Mediana	309
Simmetria	1.003313
Curtosi	0.3487307
Frequenza assoluta	744
Massimo	1105
Minimo	56.8
Deviazione standard	228.4234
Coefficiente di variazione	0.6211362
Varianza	0
Q1	213
Q3	478

Tabella 3.7: Statistiche descrittive per la variabile: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

La tabella sottolinea l'urgenza di affrontare l'inquinamento atmosferico come un problema di salute pubblica globale. La vasta gamma di mortalità indica che, mentre alcuni paesi sono riusciti a controllare efficacemente l'impatto dell'inquinamento da Particulate Matter 2.5 (PM2.5), altri sono ancora gravemente colpiti. Questi risultati evidenziano la necessità di politiche ambientali efficaci e di un maggiore impegno internazionale per ridurre l'esposizione a Particulate Matter 2.5 (PM2.5) e migliorare la salute pubblica a livello globale.

Il **boxplot** 3.49 mostra la distribuzione dei dati relativi alla mortalità da esposizione a Particulate Matter 2.5 (PM2.5). Il range di mortalità varia significativamente, con un minimo di 56.8 e un massimo di 1105 morti per milione di abitanti. Questa ampia variazione indica disparità significative tra le regioni o i paesi analizzati. La mediana è di 309, mentre la media è leggermente superiore, a 368, suggerendo una distribuzione asimmetrica con una coda più lunga verso i valori più alti. Il Primo Quartile (Q1) è a 213 e il Terzo Quartile (Q3) a 478, mostrando che la metà centrale dei dati si concentra in questo intervallo.

Un **coefficiente di simmetria** di 1.003313 indica una distribuzione asimmetrica positiva, con una coda che si estende verso i valori più alti. Questo suggerisce che alcuni paesi hanno livelli di mortalità eccezionalmente elevati. La **curtosi** di 0.3487307 suggerisce una distribuzione leggermente più appuntita rispetto a una distribuzione normale, indicando che ci sono alcuni paesi con valori estremi di mortalità. Una **deviazione standard** di 228.4234 mostra una variabilità significativa nella mortalità tra i paesi, sottolineando le differenze nelle condizioni ambientali e nelle risposte

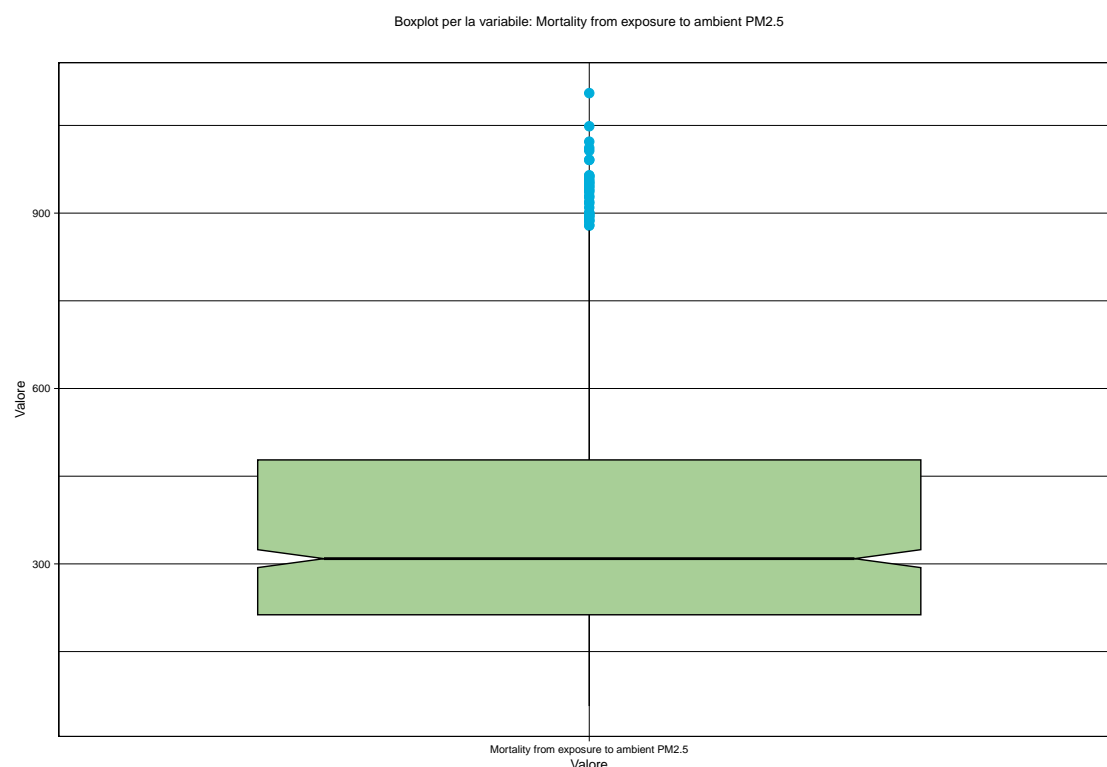


Figura 3.49: Boxplot per la variabile: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

politiche. Il **coefficiente di variazione** di 0.6211362 indica che c'è una variabilità sostanziale rispetto alla media, confermando la diversità delle condizioni tra i paesi.

L'**istogramma** 3.50 fornisce una rappresentazione visiva della frequenza dei dati relativi alla mortalità da esposizione a Particulate Matter 2.5 (PM2.5). La maggior parte dei dati si concentra nei valori inferiori, con un numero decrescente di paesi che presentano valori più alti di mortalità. L'istogramma mostra un'evidente asimmetria positiva, con più paesi che hanno una mortalità più bassa da esposizione a Particulate Matter 2.5 (PM2.5) e meno paesi con valori estremamente alti.

La **funzione di densità** offre una rappresentazione continua della distribuzione dei dati. La curva mostra una netta asimmetria positiva, confermando l'osservazione fatta sull'istogramma. Con una curtosi di 0.3487307, la distribuzione mostra un picco meno pronunciato rispetto a una distribuzione normale, suggerendo una maggiore diffusione dei valori medio-alti.

Questi grafici e la tabella indicano che, mentre un certo numero di paesi ha valori relativamente bassi di mortalità da esposizione a Particulate Matter 2.5 (PM2.5), esiste una significativa minoranza con valori molto più alti. Questo potrebbe riflettere differenze nelle condizioni ambientali, politiche di controllo dell'inquinamento, pratiche industriali e urbane, nonché nell'accesso alle cure sanitarie. La presenza di valori estremamente alti sottolinea la gravità degli impatti sulla salute dell'inquinamento atmosferico in alcune regioni. Questi risultati evidenziano la necessità urgente di interventi mirati per ridurre l'esposizione a Particulate Matter 2.5 (PM2.5) e migliorare la salute pubblica, in particolare nei paesi o nelle aree con i livelli più elevati di mortalità correlata.

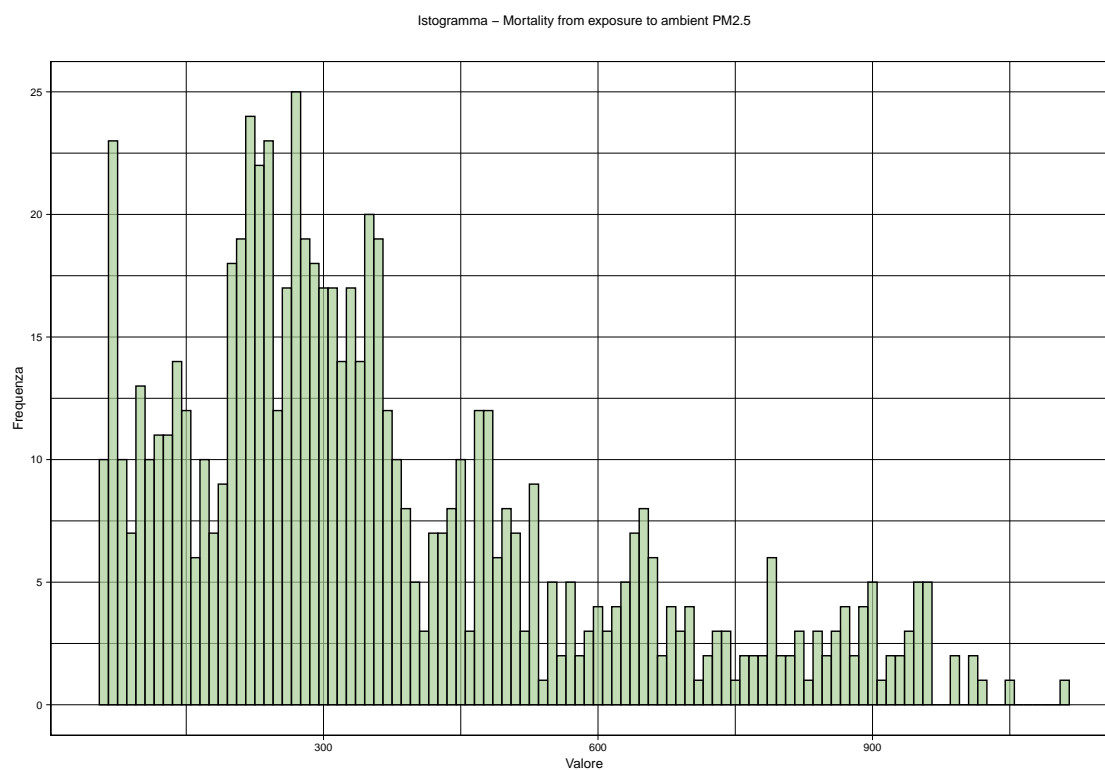


Figura 3.50: Istogramma per la variabile: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

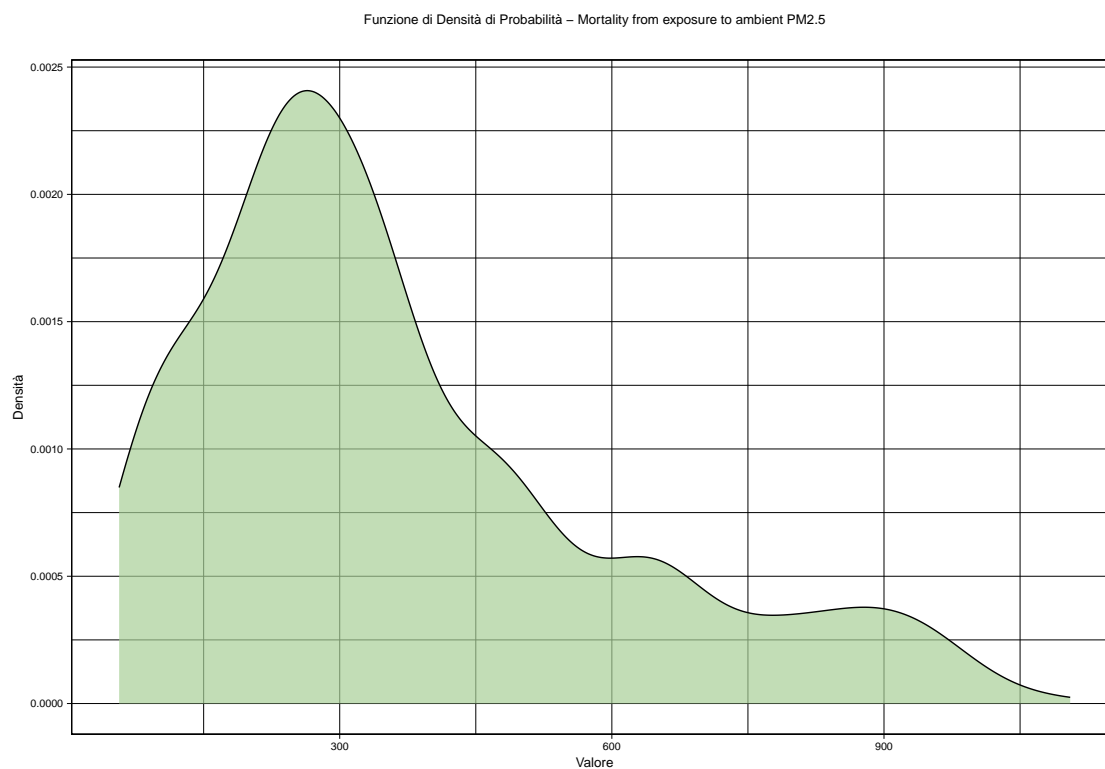


Figura 3.51: Funzione di Densità per la variabile: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

La figura 3.52 che riguarda i **paesi delle Americhe** mostra andamenti diversificati nella mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5), evidenziando come differenti nazioni affrontino e rispondano a questo problema ambientale e di salute pubblica.

Gli **Stati Uniti** mostrano una rapida e marcata diminuzione della mortalità da esposizione a Particulate Matter 2.5 (PM2.5), passando da 380 a 150 morti per milione di abitanti. Il **Canada** mostra una diminuzione costante della mortalità da esposizione a Particulate Matter 2.5 (PM2.5), passando da circa 200 a 100 morti per milione di abitanti. Il **Brasile** mantiene una mortalità relativamente stabile intorno ai 200 decessi per milione di abitanti, sebbene con periodi di picchi. L'**Argentina** mostra una tendenza generale alla diminuzione, passando da circa 400 a meno di 300 morti per milione di abitanti, nonostante alcuni periodi di crescita e decrescita. Nonostante i progressi complessivi, l'Argentina potrebbe affrontare sfide nel mantenere una riduzione costante dei livelli di Particulate Matter 2.5 (PM2.5). Il **Messico** segue un andamento simile all'Argentina, con un trend di riduzione partendo da un livello più alto. Questo potrebbe suggerire che il Messico, sebbene abbia affrontato sfide maggiori in termini di inquinamento atmosferico, sta compiendo passi significativi verso la riduzione dell'impatto del Particulate Matter 2.5 (PM2.5) sulla salute pubblica.

In generale, i paesi delle Americhe mostrano un impegno nel ridurre la mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5), sebbene con ritmi e modalità differenti. Mentre alcuni paesi come gli Stati Uniti e il Canada mostrano notevoli progressi, altri come il Brasile, l'Argentina e il Messico stanno lavorando per affrontare le proprie sfide uniche. Questi risultati evidenziano l'importanza di strategie ambientali e di salute pubblica mirate e personalizzate per affrontare efficacemente l'inquinamento atmosferico nelle diverse regioni.

La figura 3.53 illustra le tendenze della mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5) in diversi **paesi asiatici**, evidenziando pattern variabili e cambiamenti significativi nel corso del tempo.

In **Cina**, si osserva un aumento sostenuto nella mortalità da esposizione a Particulate Matter 2.5 (PM2.5), passando da 500 a quasi 1000 decessi per milione di abitanti, senza mostrare segni di decrescita. Questa tendenza all'aumento potrebbe essere dovuta a un'intensa attività industriale e all'urbanizzazione rapida, che hanno probabilmente contribuito all'aumento dei livelli di inquinamento. La dipendenza della Cina da fonti energetiche ad alta intensità di carbonio, come il carbone, potrebbe avere un impatto diretto sulla qualità dell'aria. Anche l'**India** mostra un aumento simile, con i decessi per milione di abitanti che raddoppiano da 300 a 600. Questo aumento potrebbe essere influenzato dalla rapida crescita economica e dall'aumento della popolazione urbana, che potrebbero aumentare la domanda di energia e di conseguenza peggiorare l'inquinamento atmosferico.

La **Russia** mostra un andamento altalenante, con una tendenza generale alla diminuzione, passando da circa 1000 a 500 decessi per milione di abitanti. **Giappone e Corea del Sud** mostrano andamenti simili, ma con livelli di mortalità diversi. Il Giappone passa da 200 a 300, mentre la Corea del Sud si mantiene stabile intorno ai 400 decessi per milione di abitanti. Queste tendenze potrebbero indicare che entrambi i paesi hanno sistemi avanzati di gestione ambientale, ma possono affrontare sfide diverse nella riduzione dell'esposizione a Particulate Matter 2.5 (PM2.5).

L'**Indonesia** mostra un aumento significativo, passando da 200 a quasi 400 decessi per milione di

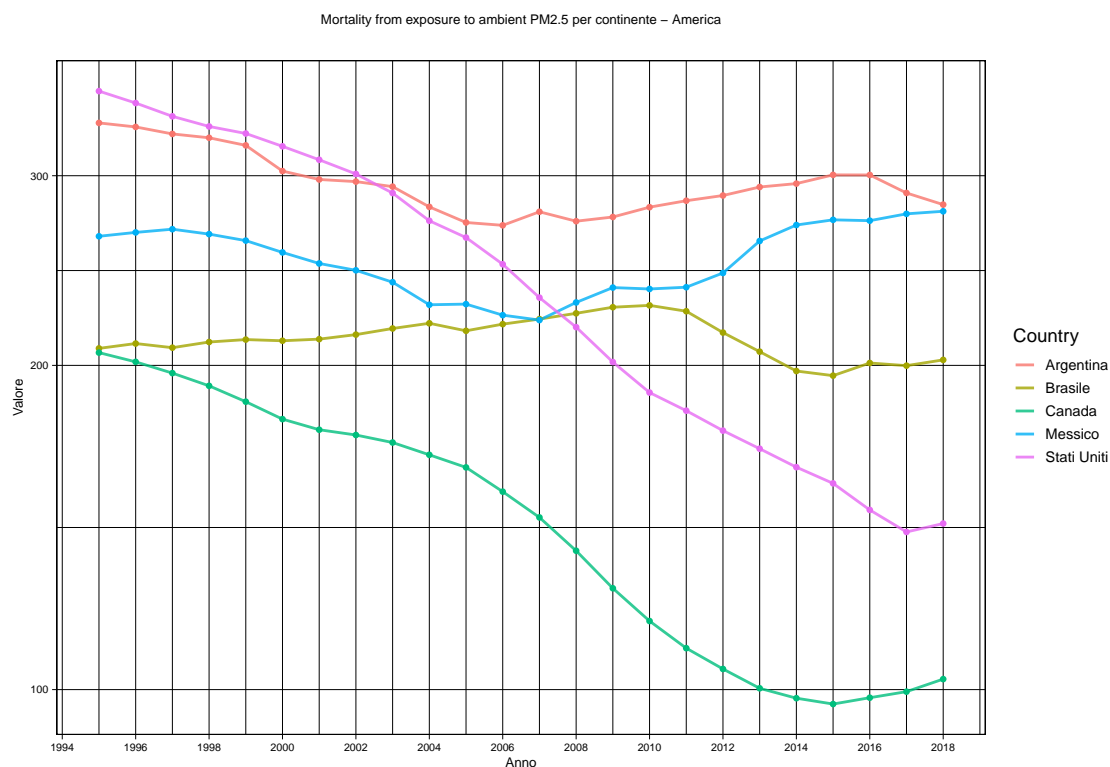


Figura 3.52: Serie storica per il continente Americano: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

abitanti. Questo aumento potrebbe essere legato allo sviluppo urbano e industriale in Indonesia, che potrebbe contribuire all'aumento dei livelli di Particulate Matter 2.5 (PM2.5).

La figura 3.54 presenta un'analisi interessante sulla mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5) in vari paesi dell'**Europa Occidentale**, mettendo in luce differenti traiettorie e andamenti nel tempo.

L'**Italia** mostra una significativa riduzione della mortalità da esposizione a Particulate Matter 2.5 (PM2.5), passando da circa 700 a 400 decessi per milione di abitanti. Il Belgio mostra una tendenza generale alla diminuzione, passando da 600 a 300 decessi per milione di abitanti, sebbene con lievi picchi di aumento. Questi picchi potrebbero indicare che, nonostante il trend generale positivo, esistono ancora sfide da affrontare nella gestione della qualità dell'aria.

I **Paesi Bassi** partono da circa 550 decessi per milione di abitanti e raggiungono circa 300, con alcuni lievi aumenti nel corso del tempo. Questo andamento indica che, pur essendoci un progresso complessivo, la riduzione dell'inquinamento e delle sue conseguenze sulla salute è un processo con sfide e fluttuazioni.

Il **Lussemburgo** mostra una riduzione drastica, passando da circa 450 a meno di 100 decessi per milione di abitanti. Questo calo notevole potrebbe riflettere la scala ridotta del paese, che può facilitare l'attuazione di politiche efficaci e interventi rapidi.

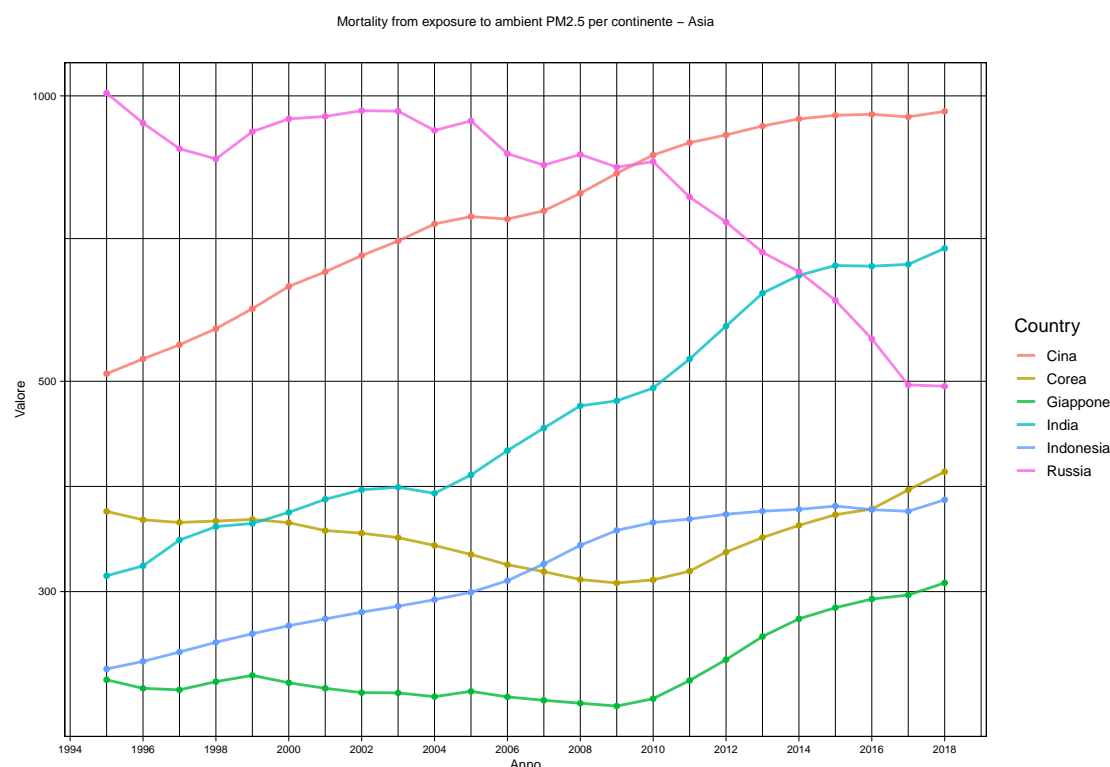


Figura 3.53: Serie storica per il continente Asiatico: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

Il **Regno Unito** ha registrato una rapida diminuzione della mortalità, passando da più di 600 a 440 decessi per milione di abitanti. Questo potrebbe essere il risultato di politiche ambientali e sanitarie efficaci volte a migliorare la qualità dell'aria e a ridurre l'esposizione ai Particulate Matter 2.5 (PM2.5).

In generale, i paesi dell'Europa Occidentale sembrano seguire un trend di diminuzione della mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5), sebbene con velocità e andamenti diversi. Queste tendenze indicano che, nonostante le sfide persistenti, ci sono stati progressi significativi nella riduzione dell'impatto dell'inquinamento atmosferico sulla salute. Questi risultati sottolineano l'importanza di continuare a investire in politiche ambientali efficaci e in tecnologie pulite per mantenere e accelerare questi miglioramenti.

La figura 3.55 fornisce un'analisi approfondita della mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5) in vari paesi dell'**Europa Orientale**. Dall'analisi dei dati emerge un panorama differenziato, con ciascun paese che mostra tendenze uniche nella lotta contro l'inquinamento atmosferico e i suoi effetti sulla salute.

Finlandia, Svezia e Norvegia mostrano un modello di mortalità molto simile, caratterizzato da una tendenza decrescente. Partendo da valori tra i 250 e i 300 decessi per milione di abitanti, hanno registrato una diminuzione costante, con periodi di stallo ma senza rialzi significativi. L'impegno di questi paesi per l'ambiente e per la qualità dell'aria potrebbe aver giocato un ruolo chiave nella riduzione della mortalità.

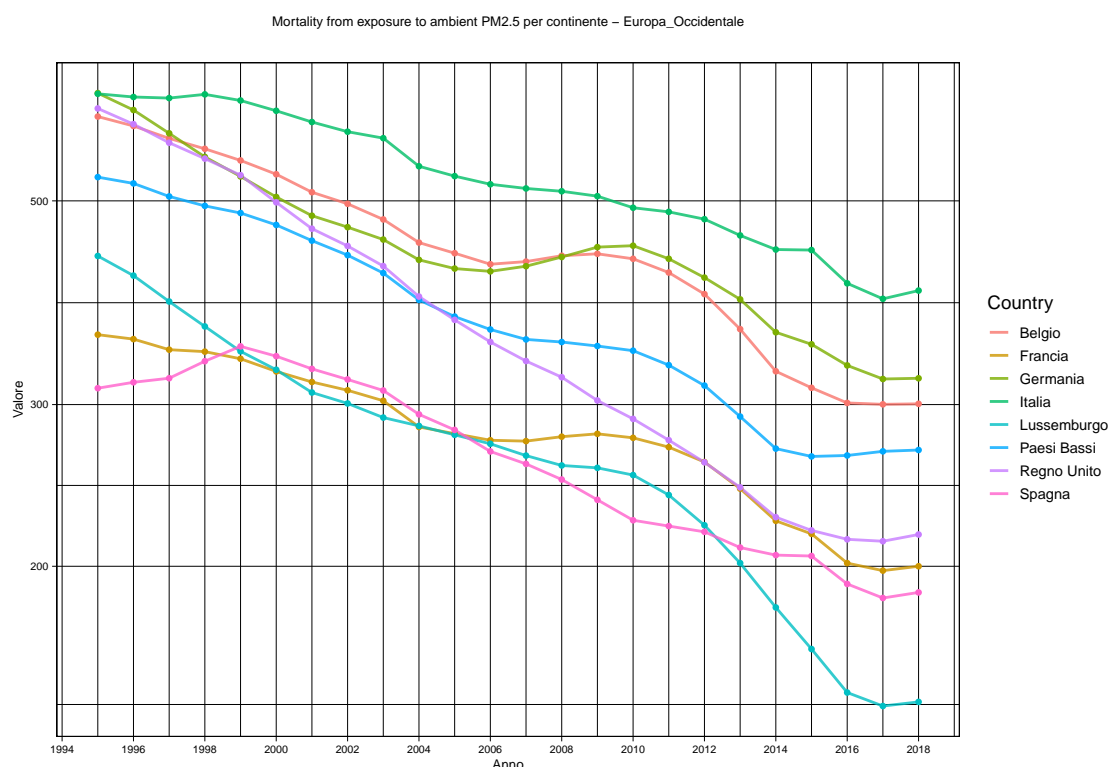


Figura 3.54: Serie storica per i paesi dell'Europa Occidentale: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

L'**Ungheria** presenta il numero più elevato di decessi, ma anche qui si nota una tendenza alla diminuzione, da oltre 900 a circa 700 morti per milione di abitanti. La **Polonia**, partendo da un livello leggermente inferiore rispetto all'Ungheria, mostra un trend simile e raggiunge un numero di decessi paragonabile. Questa tendenza potrebbe indicare sfide comuni nell'Europa Orientale nella gestione della qualità dell'aria.

La **Grecia** mostra un andamento più stabile ma con una tendenza generale alla diminuzione, il che potrebbe riflettere una combinazione di fattori, tra cui il clima, le politiche ambientali e le pratiche industriali.

La **Danimarca** ha registrato la diminuzione più marcata, partendo da circa 600 decessi per milione di abitanti e arrivando a 250. Questo calo rapido potrebbe essere il risultato di un impegno deciso e di politiche efficaci nell'ambito della qualità dell'aria e della salute pubblica.

La figura 3.56 offre un'analisi interessante sulle tendenze della mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5) nei paesi dell'**Oceania** e dell'**Africa**, evidenziando differenze significative nei modelli di mortalità tra queste regioni.

In **Oceania**, **Nuova Zelanda** e **Australia** mostrano un andamento decrescente nella mortalità da esposizione a Particulate Matter 2.5 (PM2.5). Entrambi i paesi potrebbero aver implementato politiche ambientali efficaci e misure di controllo dell'inquinamento per ridurre l'esposizione a

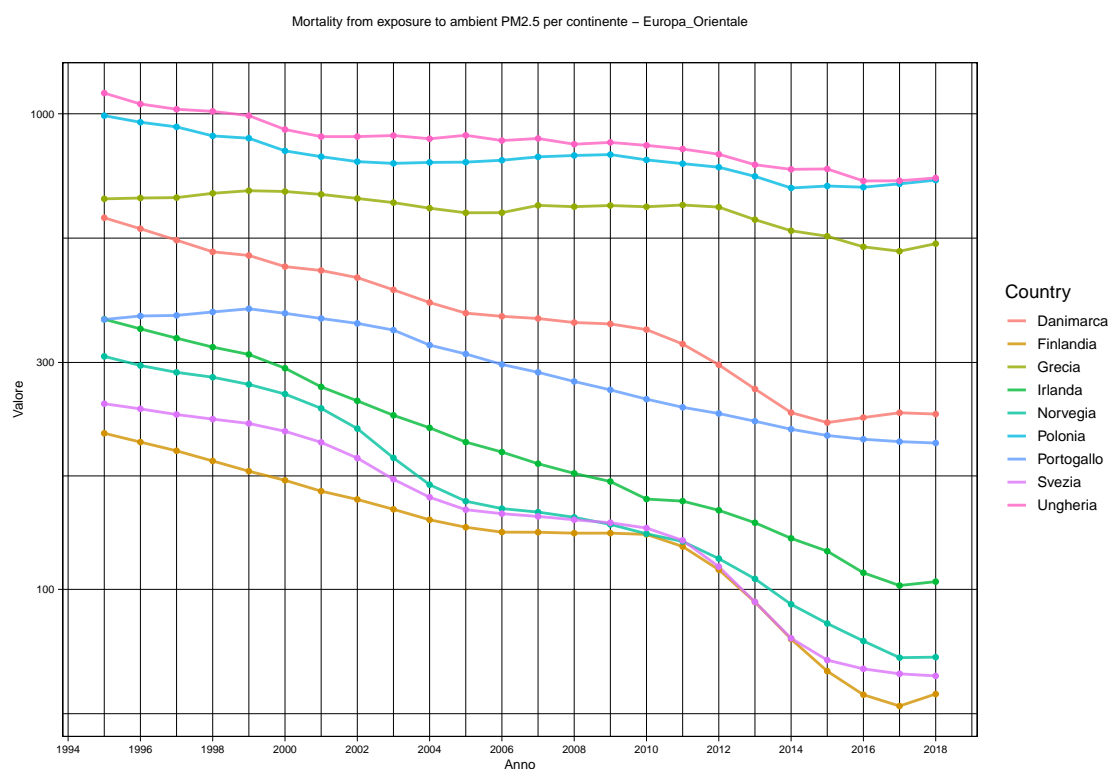


Figura 3.55: Serie storica per i paesi dell'Europa Orientale: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

Particulate Matter 2.5 (PM2.5).

Il **Sud Africa**, invece, mostra un andamento molto più variabile con un picco significativo nel 2009 e nel 2010, seguito da una diminuzione, ma mantenendosi comunque a livelli più elevati rispetto all'inizio della serie storica. Il picco osservato potrebbe essere il risultato di variazioni nelle politiche ambientali, cambiamenti economici o eventi specifici che hanno portato a un aumento temporaneo dell'inquinamento atmosferico. Come paese in via di sviluppo, il Sud Africa potrebbe affrontare sfide maggiori nell'equilibrare crescita economica e protezione ambientale. Una maggiore dipendenza da fonti energetiche inquinanti, come il carbone, potrebbe spiegare i livelli più elevati di Particulate Matter 2.5 (PM2.5).

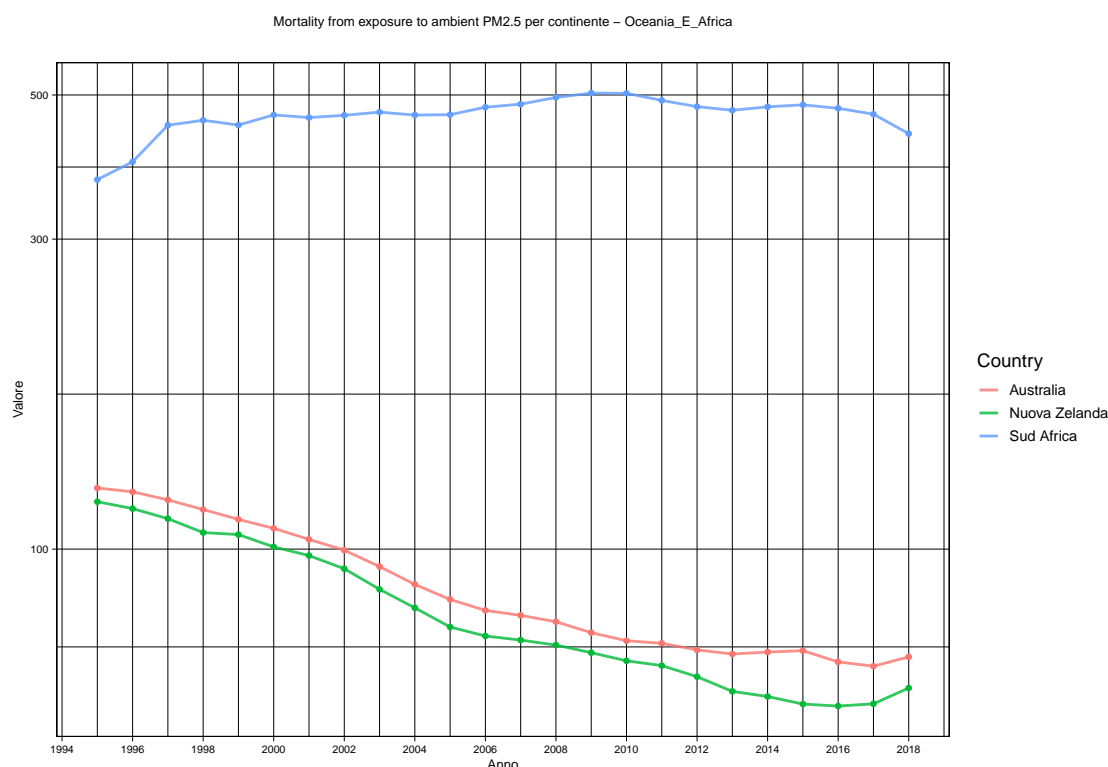


Figura 3.56: Serie storica per i paesi dell'Oceania e dell'Africa: Mortalità da esposizione a Particulate Matter 2.5 (PM2.5)

3.2.8 Percentuale di popolazione esposta a più di $10 \mu\text{g}/\text{m}^3$

Questo indicatore rappresenta la percentuale di popolazione esposta a concentrazioni annuali di particolato PM2,5 superiori al valore raccomandato dalla World Health Organization (WHO) per la qualità dell'aria, ovvero 10 microgrammi per metro cubo ($10 \mu\text{g}/\text{m}^3$).

I dati sull'esposizione all'inquinamento atmosferico sono ottenuti dal dataset "Exposure to Particulate Matter 2.5 (PM2.5) – countries and regions" proveniente dall' Database sull'Ambiente dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE).

Ulteriori dettagli sulla metodologia utilizzata per calcolare l'esposizione all'inquinamento atmosferico e le fonti di dati sottostanti sono disponibili nella relativa metadata (metadati associati).

La **media** è del 74.2%, mentre la **mediana** è significativamente più alta, a 96.4%. Questa discrepanza tra media e mediana suggerisce una distribuzione asimmetrica dei dati. Un **coefficiente di simmetria** di -1.159619 indica una distribuzione asimmetrica negativamente, con più paesi che hanno una percentuale bassa di popolazione esposta. La **curtosi** di -0.323619 suggerisce una distribuzione leggermente meno appuntita rispetto a una normale. Un'ampia variazione è evidente tra il valore minimo (0%) e il massimo (100%), indicando significative differenze nella qualità dell'aria tra i paesi. La **deviazione standard** di 36.43437 e la varianza di 1327.463 riflettono una notevole variabilità tra i paesi. Q1 a 56.2% e Q3 a 99.5% mostrano una distribuzione concentrata verso valori più alti, con la metà dei paesi che ha una percentuale di esposizione superiore al 56.2%.

Il **boxplot** 3.57 mostra una mediana alta e una distribuzione asimmetrica verso l'alto. La presenza

Percentuale di popolazione esposta a più di 10 $\mu\text{g}/\text{m}^3$	
Media	74.2
Mediana	96.4
Simmetria	-1.159619
Curtosi	-0.323619
Frequenza assoluta	744
Massimo	100
Minimo	0
Deviazione standard	36.43437
Coefficiente di variazione	0.4910036
Varianza	1327.463
Q1	56.2
Q3	99.5

Tabella 3.8: Statistiche descrittive per la variabile: Percentuale di popolazione esposta a più di 10 $\mu\text{g}/\text{m}^3$

di valori estremamente bassi è indicata dalla lunga coda verso il basso.

L'**istogramma** 3.58) evidenzia una prevalenza di valori elevati di esposizione a Particulate Matter 2.5 (PM2.5), con una concentrazione significativa di paesi che hanno alte percentuali di popolazione esposta.

La **funzione di densità** conferma la simmetria negativa osservata, con un picco nei valori più alti e una coda lunga verso i valori più bassi.

Questi dati suggeriscono che, nonostante la crescente consapevolezza e gli sforzi per ridurre l'inquinamento atmosferico, una significativa percentuale della popolazione mondiale rimane esposta a livelli di Particulate Matter 2.5 (PM2.5) che superano le linee guida dell'Organizzazione Mondiale della Sanità (OMS). La varietà nei livelli di esposizione riflette differenze nelle politiche ambientali, nelle pratiche industriali, e nelle condizioni geografiche e climatiche tra i paesi. L'analisi evidenzia la necessità urgente di interventi mirati e politiche efficaci per migliorare la qualità dell'aria e ridurre i rischi per la salute pubblica a livello globale.

La figura 3.60 fornisce un quadro chiaro dell'andamento della percentuale di popolazione esposta a livelli di Particulate Matter 2.5 (PM2.5) superiori a 10 $\mu\text{g}/\text{m}^3$ nei paesi del **continente americano**. Ogni paese mostra un trend unico, riflettendo le diverse realtà ambientali e le politiche adottate per gestire la qualità dell'aria.

Il **Canada** mostra un trend molto positivo, con una riduzione da circa 50% a quasi 0%. La presenza di "stalli" nel grafico potrebbe indicare periodi in cui non ci sono stati miglioramenti significativi, ma la tendenza generale è chiaramente verso una diminuzione della popolazione esposta a Particulate Matter 2.5 (PM2.5). Questo può essere il risultato di efficaci politiche ambientali e di un impegno costante nel migliorare la qualità dell'aria. Gli Stati Uniti mostrano una rapida diminuzione della percentuale di popolazione esposta, passando dal 75% a circa il 5%. Questo rapido calo suggerisce un impegno significativo e successo nelle misure di controllo dell'inquinamento e nella gestione della qualità dell'aria.

Il **Messico**, al contrario, mostra una situazione statica, con la percentuale di popolazione esposta

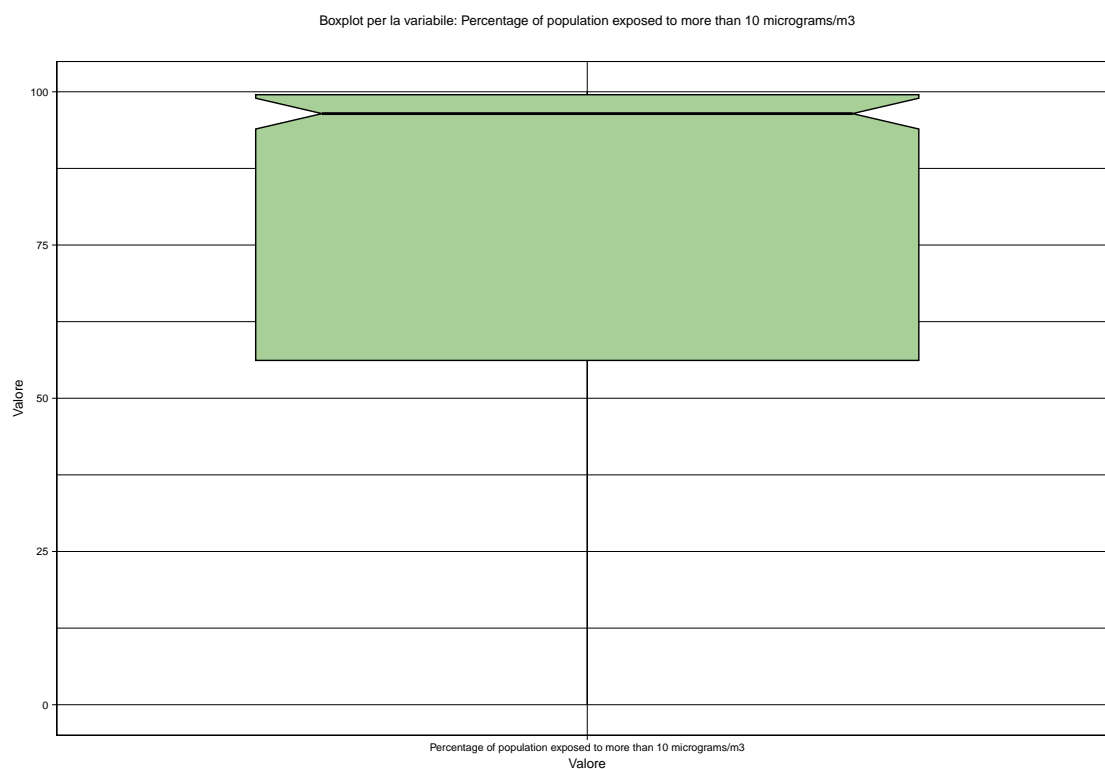
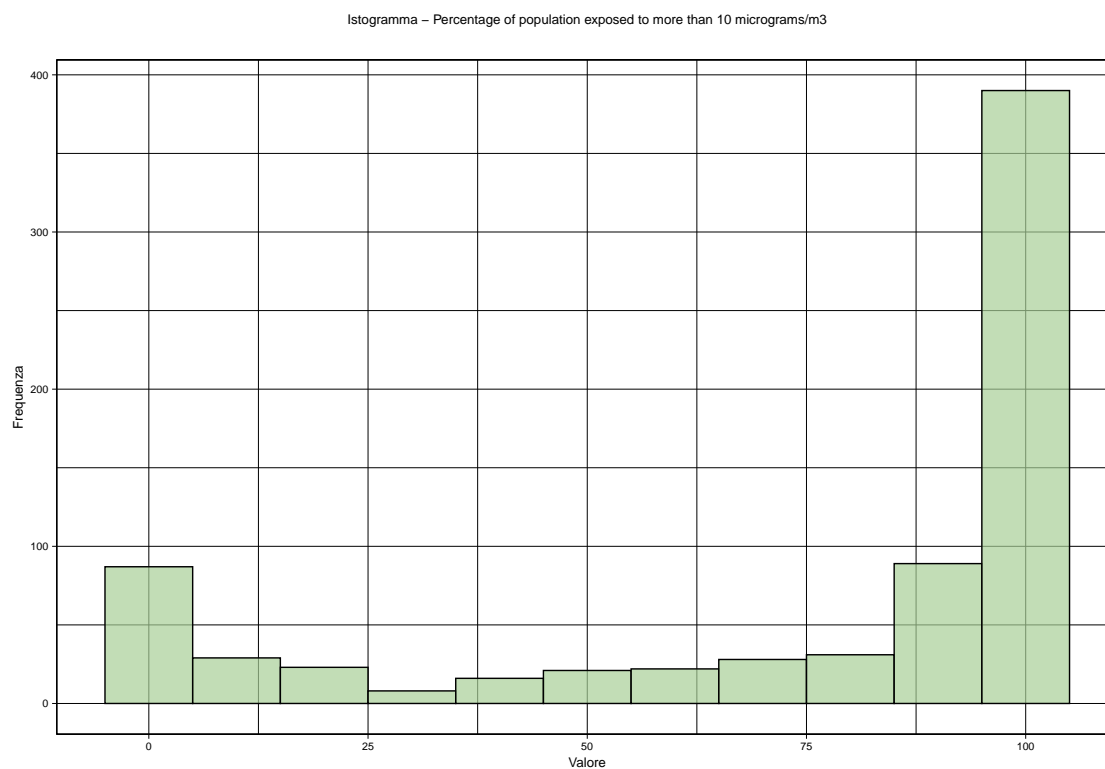
Figura 3.57: Boxplot per la variabile: Percentuale di popolazione esposta a più di 10 $\mu\text{g}/\text{m}^3$ 

Figura 3.58: Istogramma

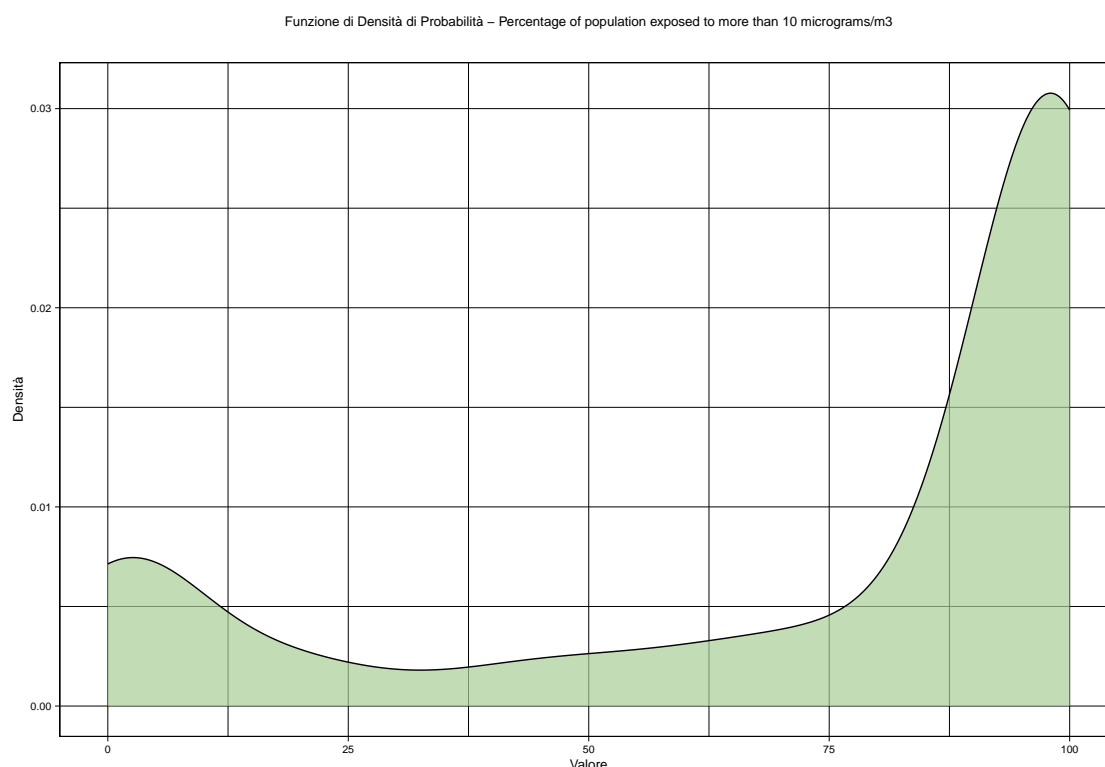


Figura 3.59: Funzione di densità

che rimane fissa al 100%. Questo indica che la maggior parte della popolazione continua a vivere in condizioni di elevata esposizione a Particulate Matter 2.5 (PM2.5), suggerendo la necessità di interventi più efficaci e di un rinnovato focus sulle politiche ambientali.

Il **Brasile** mostra un leggero calo della percentuale di popolazione esposta, passando dal 100% al 75%, con periodi di crescita e decrescita. Questa variabilità potrebbe riflettere la complessità delle sfide ambientali in un paese di grandi dimensioni e diverse condizioni geografiche e meteorologiche.

L'**Argentina** mostra una percentuale costantemente alta di esposizione, vicina al 100%, con alcuni periodi di leggera diminuzione. Come nel caso del Messico, questo suggerisce che vi sono ancora importanti sfide da affrontare per ridurre l'esposizione a Particulate Matter 2.5 (PM2.5).

La figura 3.61 mostra le tendenze della percentuale di popolazione esposta a concentrazioni di Particulate Matter 2.5 (PM2.5) superiori a 10 $\mu\text{g}/\text{m}^3$ in alcuni paesi asiatici. Ogni paese presenta un andamento distinto, riflettendo le diverse sfide ambientali e le politiche adottate per affrontare l'inquinamento atmosferico.

La **Russia** mostra una riduzione generale della percentuale di popolazione esposta, passando dal 95% al 70%. Interessante notare un cambiamento significativo nel trend dal 2014, con una diminuzione più marcata rispetto al periodo precedente. Questo potrebbe riflettere l'implementazione di politiche ambientali più efficaci o cambiamenti significativi nelle pratiche industriali e urbane.

Contrariamente, il **Giappone** mostra un aumento della percentuale di popolazione esposta, pas-

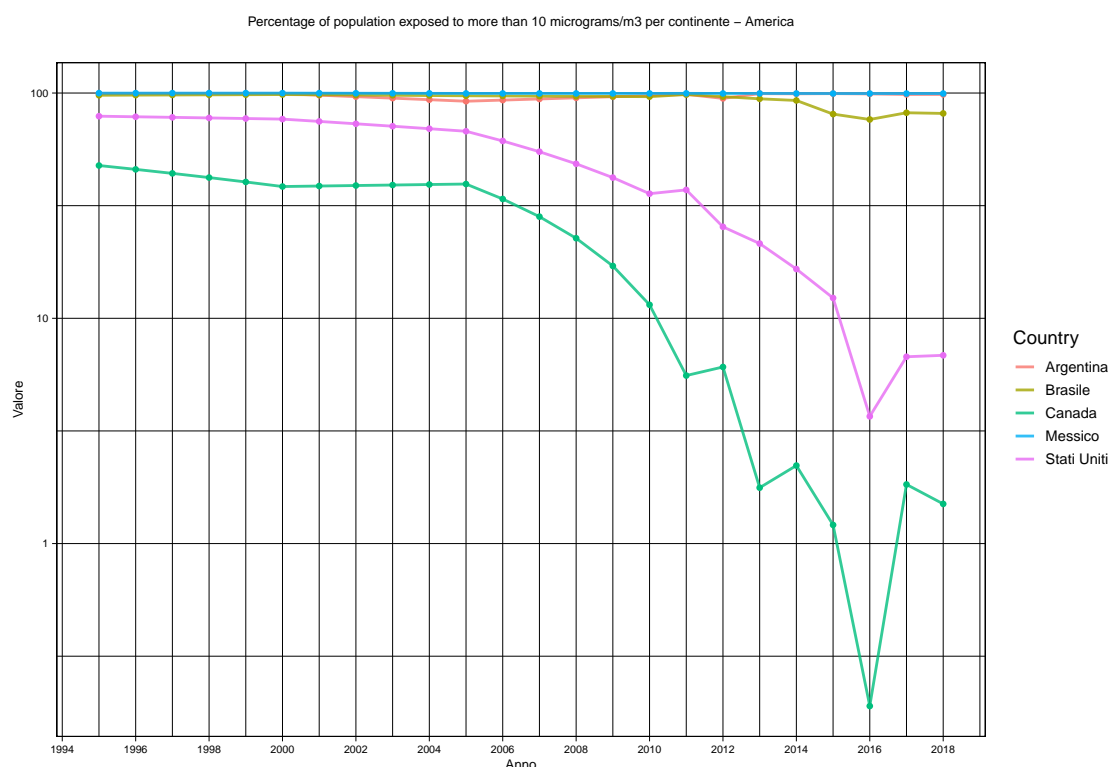


Figura 3.60: Serie storica per il continente Americano

sando dal 90% al 100%. Questo aumento potrebbe essere dovuto a vari fattori, inclusi cambiamenti nelle condizioni atmosferiche, aumento dell'urbanizzazione, o variazioni nella composizione delle emissioni di inquinanti.

Corea, Indonesia e India mostrano una situazione quasi stazionaria, con la percentuale di popolazione esposta che rimane costantemente vicina al 100%. Questo suggerisce che, nonostante eventuali sforzi per migliorare la qualità dell'aria, la maggior parte della popolazione in questi paesi continua a vivere in condizioni di elevata esposizione a Particulate Matter 2.5 (PM2.5). Questo potrebbe essere il risultato di sfide persistenti, come l'elevata densità urbana, la dipendenza da fonti energetiche inquinanti, e limitazioni nelle politiche e nelle tecnologie di controllo dell'inquinamento.

La figura 3.62 illustra l'andamento della percentuale di popolazione esposta a Particulate Matter 2.5 (PM2.5) superiore a $10 \mu\text{g}/\text{m}^3$ nei principali paesi dell'**Europa Occidentale**, offrendo una visione chiara di come queste nazioni stiano affrontando la sfida dell'inquinamento atmosferico.

La percentuale di popolazione esposta nei **Paesi Bassi** rimane costantemente alta, intorno al 100%, con solo lievi variazioni. Questo suggerisce che, nonostante gli sforzi, l'esposizione alla qualità dell'aria inferiore continua ad essere un problema significativo, probabilmente a causa della densità della popolazione e delle attività industriali.

Il **Regno Unito** mostra un trend positivo, con una riduzione significativa della popolazione esposta, passando da circa il 95% al 55%. Questo calo potrebbe essere il risultato di politiche ambientali efficaci e un maggiore impegno nel migliorare la qualità dell'aria.

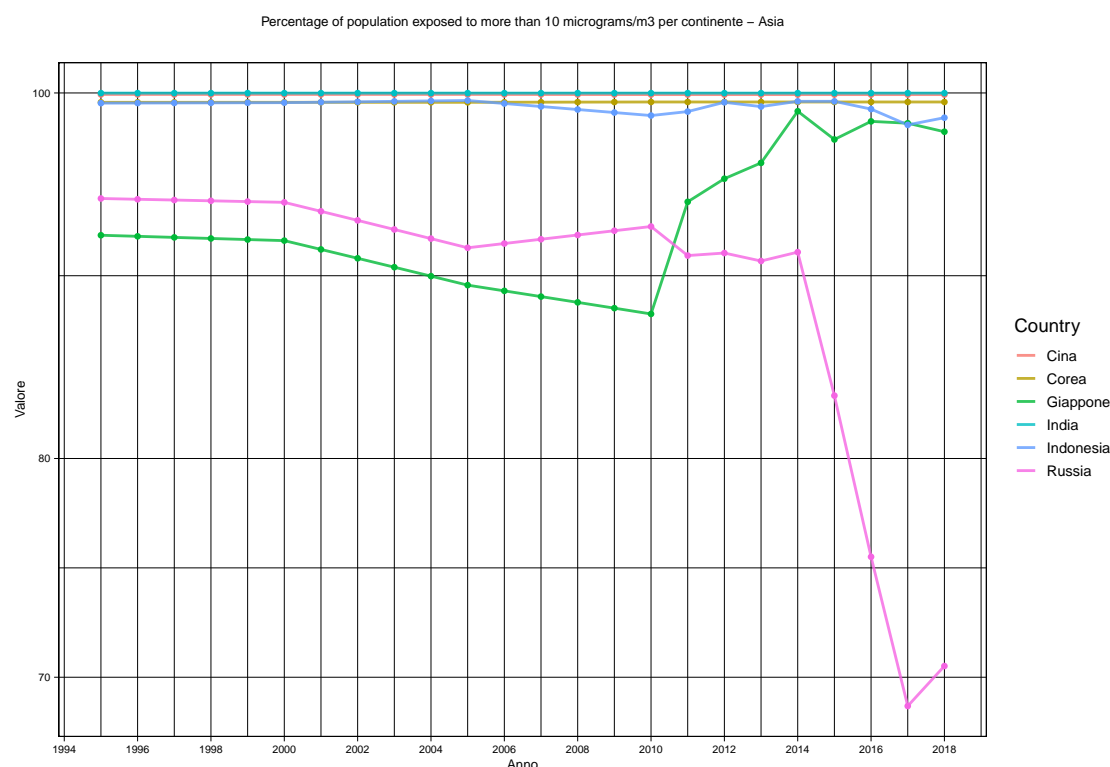


Figura 3.61: Serie storica per il continente Asiatico

La **Spagna** presenta un andamento altalenante, oscillando tra l'80% e il 45%, ma con picchi che raggiungono anche il 90%. Questa variabilità potrebbe essere attribuita a fattori stagionali, variazioni nelle politiche ambientali, o a specifici eventi che hanno influenzato temporaneamente la qualità dell'aria. Questi paesi mostrano una tendenza generale alla diminuzione, partendo dal 100% e riducendosi intorno ai 90%. La diminuzione suggerisce progressi nella gestione dell'inquinamento atmosferico, ma indica anche che c'è ancora spazio per ulteriori miglioramenti.

Il **Lussemburgo** mostra una notevole riduzione iniziale, passando dal 100% al 40%, ma poi risale al 70%. Questo andamento potrebbe riflettere cambiamenti nelle politiche locali o variazioni nelle condizioni ambientali che hanno influenzato l'esposizione a Particulate Matter 2.5 (PM2.5).

La **Francia** passa dal 90% al 70% con periodi di aumento e diminuzione. Questo potrebbe indicare una certa incostanza nelle politiche ambientali o nelle condizioni atmosferiche.

La figura 3.63 fornisce un quadro complesso e differenziato della situazione dell'esposizione a Particulate Matter 2.5 (PM2.5) in **Europa Orientale**. Esaminiamo i diversi andamenti osservati nei vari paesi.

L'**Ungheria** mostra una costante esposizione al 100%, indicando che quasi tutta la popolazione è esposta a livelli di Particulate Matter 2.5 (PM2.5) superiori ai limiti raccomandati. Questa situazione stabile suggerisce sfide persistenti nel ridurre l'inquinamento atmosferico e potrebbe riflettere la necessità di riforme significative nelle politiche ambientali.

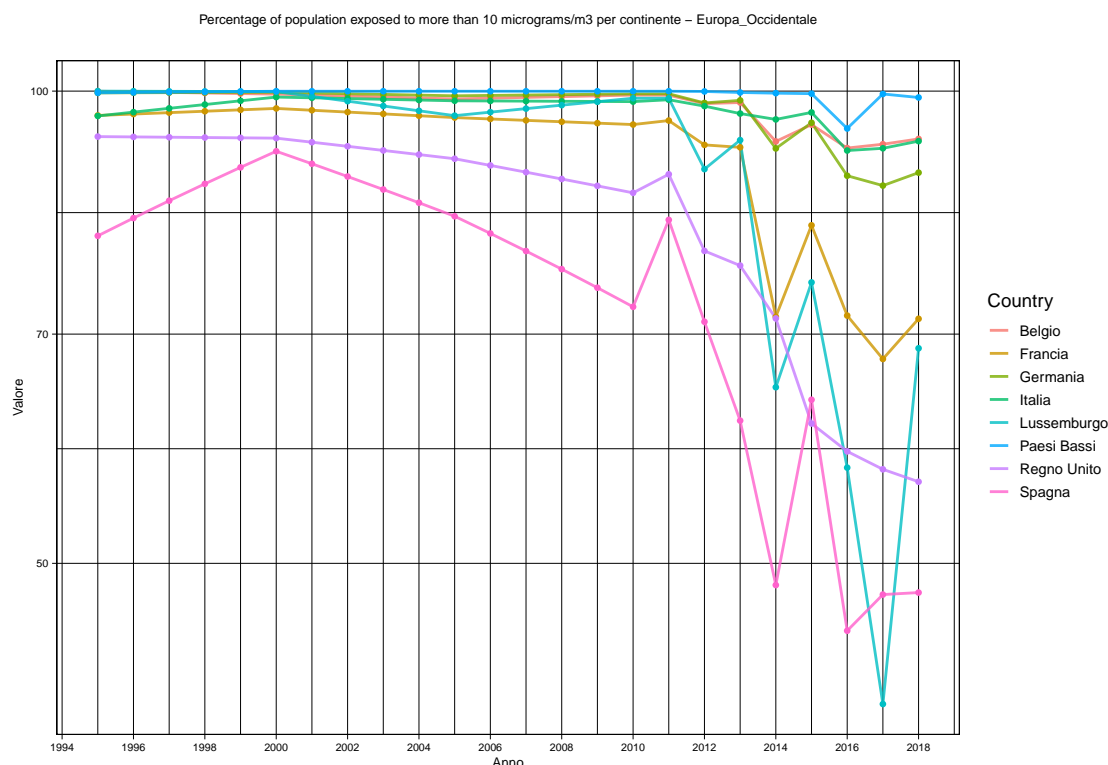


Figura 3.62: Serie storica per i paesi dell'Europa Occidentale

La **Grecia**, con un andamento non lineare ma attorno al 100%, affronta sfide simili all'Ungheria. La consistente alta esposizione potrebbe essere attribuita a fattori come la densità urbana, le condizioni meteorologiche e le emissioni industriali.

La **Danimarca** presenta un'interessante variazione, con una diminuzione iniziale da circa il 95% al 25% nel 2016, seguita da un lieve incremento al 30% nel 2018. Questo trend indica un miglioramento significativo nella gestione dell'inquinamento atmosferico, sebbene con alcune fluttuazioni.

L'**Irlanda** mostra un marcato miglioramento, con una riduzione da circa l'80% a zero. Nonostante alcuni periodi di crescita, la tendenza generale alla diminuzione suggerisce un impegno efficace nel ridurre l'esposizione a Particulate Matter 2.5 (PM2.5).

Il **Portogallo** inizia dal 60%, con una riduzione non costante che scende al 20% nel 2016 e poi risale leggermente al 23% nel 2018. Questo andamento indica un certo progresso, ma anche la presenza di sfide nel mantenere una riduzione costante dell'esposizione. Svezia e Finlandia mostrano una rapida diminuzione dell'esposizione, con la Svezia che scende quasi a zero e la Finlandia che arriva a zero. Questi trend positivi riflettono un efficace controllo dell'inquinamento atmosferico e una forte attenzione alla qualità dell'aria.

La figura 3.64 offre una visione contrastante dell'esposizione a Particulate Matter 2.5 (PM2.5) superiore a 10 $\mu\text{g}/\text{m}^3$ in **Africa e Oceania**, riflettendo le diverse realtà ambientali e le risposte politiche in questi due continenti.

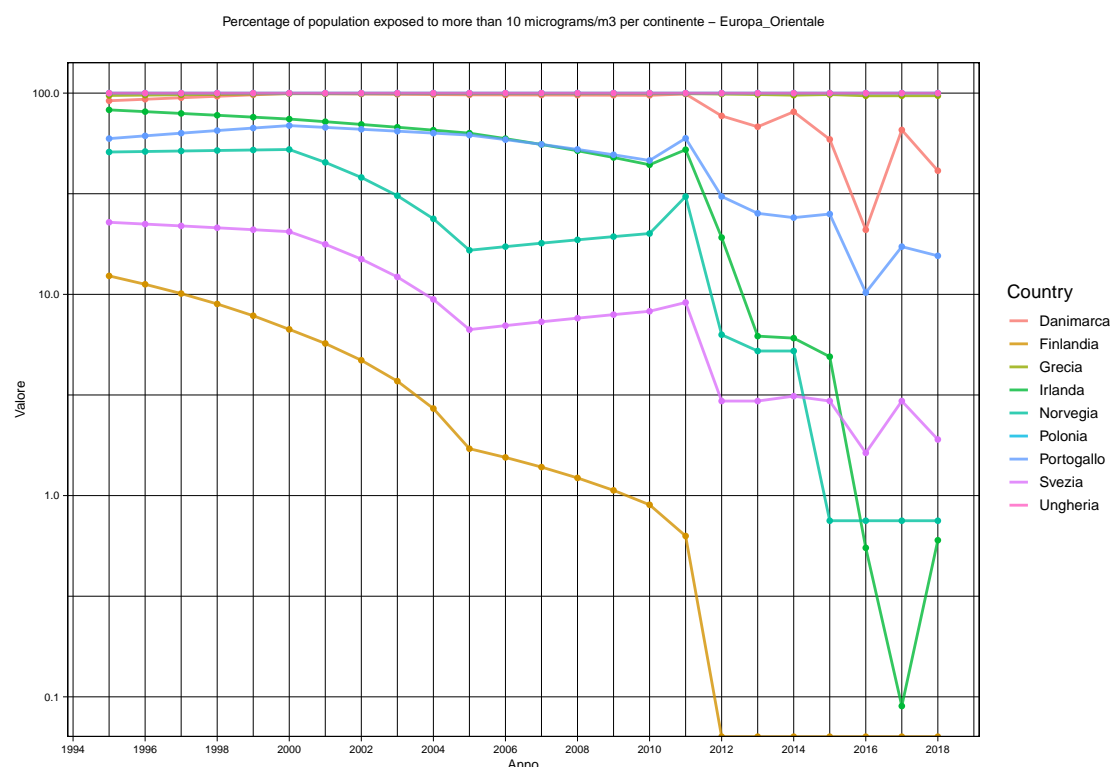


Figura 3.63: Serie storica per i paesi dell'Europa Orientale

Il **Sud Africa** mostra una situazione critica, con una costante esposizione al 100%. Questo indica che quasi tutta la popolazione è regolarmente esposta a livelli di Particulate Matter 2.5 (PM2.5) che superano i limiti raccomandati. Tale costanza suggerisce sfide persistenti nel gestire l'inquinamento atmosferico, che possono essere legate a una combinazione di fattori urbani, industriali e naturali. La mancanza di miglioramenti in questa regione sottolinea la necessità di riforme significative nelle politiche ambientali e di un impegno maggiore per ridurre l'inquinamento dell'aria.

Al contrario, **Australia e Nuova Zelanda** mostrano un quadro molto positivo, con una costante esposizione a zero. Questo implica che la popolazione in questi paesi non è generalmente esposta a livelli di Particulate Matter 2.5 (PM2.5) superiori a $10 \mu\text{g}/\text{m}^3$. Questi risultati riflettono l'efficacia delle politiche ambientali e delle strategie di gestione della qualità dell'aria implementate in questi paesi. Potrebbe anche indicare una combinazione favorevole di condizioni naturali, densità di popolazione relativamente bassa e un'industrializzazione controllata che contribuisce a mantenere bassi livelli di inquinamento atmosferico.

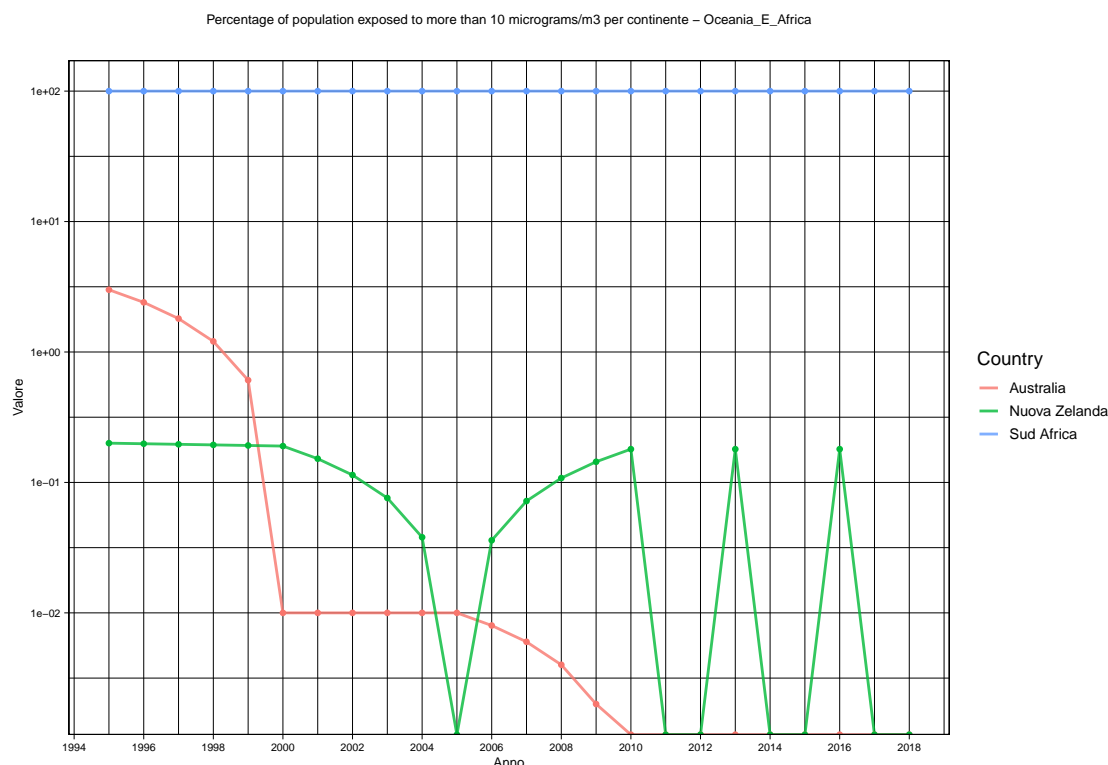


Figura 3.64: Serie storica per i paesi dell'Oceania e dell'Africa

3.2.9 Costi del Benessere per le Mortalità Precoce dovute all'Esposizione a Particulate Matter 2.5 (PM2.5) Ambientali, Equivalente Prodotto Interno Lordo (PIL)

Il costo sociale delle morti premature causate dall'esposizione a Particulate Matter 2.5 (PM2.5) ambientali, espresso in termini di equivalente del Prodotto Interno Lordo (PIL) (Prodotto Interno Lordo), rappresenta la misura dei costi derivanti dalle morti premature attribuibili a rischi ambientali, confrontati con il Prodotto Interno Lordo (PIL) come punti percentuali di equivalente del Prodotto Interno Lordo (PIL).

Le stime dei costi rappresentano solo il costo delle morti premature e sono calcolate utilizzando stime del Valore di una Vita Statistica e il numero di morti premature attribuibili a ciascun rischio ambientale. Escludono gli impatti sulla morbilità (perdite di produttività del lavoro, costi di trattamento e disposizione a pagare per evitare dolore e sofferenza dovuti a malattie). Escludono anche impatti diversi da quelli sulla salute umana (ad esempio, sugli edifici, la produttività agricola, la salute degli ecosistemi). Il costo sociale dell'esposizione a questi rischi ambientali è quindi maggiore rispetto ai costi delle morti presentati qui. Tuttavia, le evidenze disponibili suggeriscono che i costi legati alle morti rappresentino la maggior parte dei costi totali per la società. Inoltre, il Valore di una Vita Statistica (VSL) cattura anche valori non di mercato non correlati alle spese e quindi non fa parte integrante del calcolo del Prodotto Interno Lordo (PIL). Di conseguenza, le stime dei costi vengono confrontate solo a titolo illustrativo con il Prodotto Interno Lordo (PIL).

I dati e ulteriori dettagli sulla metodologia possono essere ottenuti dal Global Burden of Disease

Study 2019 Results (Studio Globale del Carico di Malattia 2019) e dal dataset sulla mortalità e i costi sociali derivanti dall'esposizione a rischi ambientali dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE)).

Costi del Benessere per le Mortalità Precoce dovute all'Esposizione a PM2.5 Ambientali	
Media	3.86
Mediana	3.14
Simmetria	1.111039
Curtosi	0.5248807
Frequenza assoluta	744
Frequenza relativa	0
Massimo	12.2
Minimo	0.56
Deviazione standard	2.578509
Coefficiente di variazione	0.6674591
Varianza	6.648711
Q1	2.18
Q3	5.12

Tabella 3.9: Statistiche descrittive per la variabile: Costi del Benessere per le Mortalità Precoce dovute all'Esposizione a Particulate Matter 2.5 (PM2.5) Ambientali

La **media** di 3.86 e la **mediana** di 3.14 indicano una distribuzione asimmetrica positiva, suggerendo che alcuni paesi subiscono costi del benessere molto più elevati rispetto alla media. L'**asimmetria positiva**, con un coefficiente di simmetria di 1.111039, e una **curtosi** di 0.5248807, evidenziano una distribuzione con una coda estesa verso valori più alti, suggerendo che alcuni paesi hanno costi del benessere eccezionalmente elevati.

La variazione significativa tra il **valore minimo** di 0.56 e il **massimo** di 12.2 riflette le grandi differenze nei costi del benessere tra i paesi. Questa variabilità, come indicato dalla deviazione standard di 2.578509, evidenzia come i costi del benessere possano variare ampiamente a seconda delle condizioni locali, delle politiche ambientali e della vulnerabilità all'inquinamento atmosferico.

Il **boxplot** 3.65 illustra questa varietà di situazioni, con una mediana più bassa e una coda lunga verso i valori più elevati. Questo indica che mentre la maggior parte dei paesi ha costi del benessere relativamente bassi, un numero significativo di paesi affronta costi molto più elevati.

L'**istogramma** 3.66 sottolinea ulteriormente la prevalenza di valori più bassi nella distribuzione, mostrando una concentrazione di paesi con costi del benessere inferiori alla media. Questo sottolinea che, sebbene l'impatto medio a livello globale sia relativamente moderato, esistono notevoli disparità tra i paesi.

La **funzione di densità** 3.67 conferma la presenza di una coda estesa verso i valori più elevati, indicando che i costi del benessere elevati, sebbene meno comuni, rappresentano una sfida significativa per alcuni paesi.

Questi dati suggeriscono che l'impatto economico dell'esposizione a Particulate Matter 2.5 (PM2.5) varia notevolmente tra i paesi. Mentre molti paesi hanno costi del benessere relativamente bassi, alcuni mostrano un impatto economico molto più elevato. Questa variabilità riflette le differenze

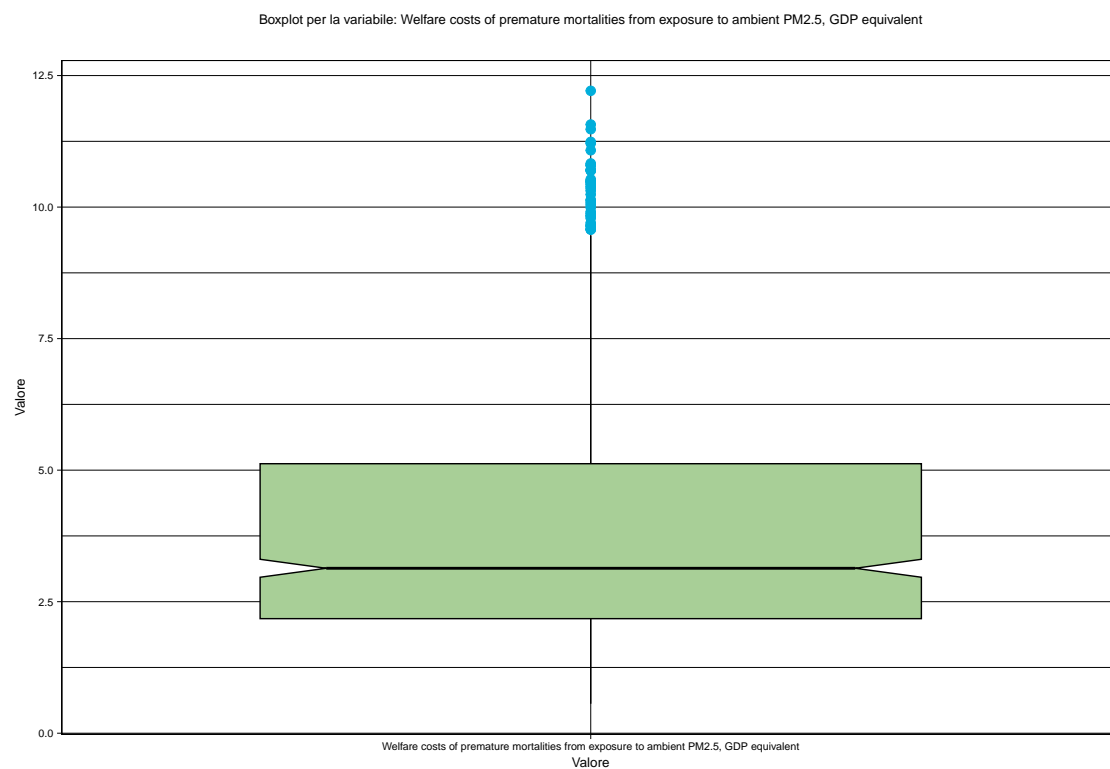


Figura 3.65: Boxplot per la variabile:Popolazione con accesso a fonti di acqua potabile migliorate, % della popolazione totale

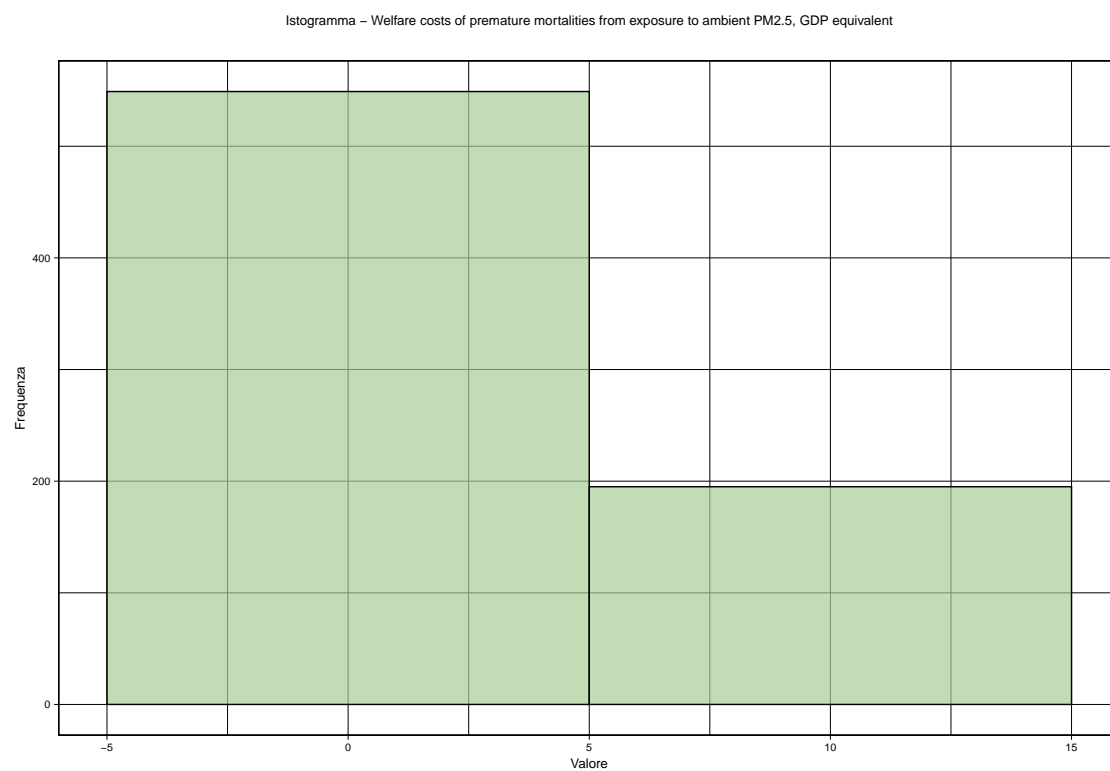


Figura 3.66: Istogramma

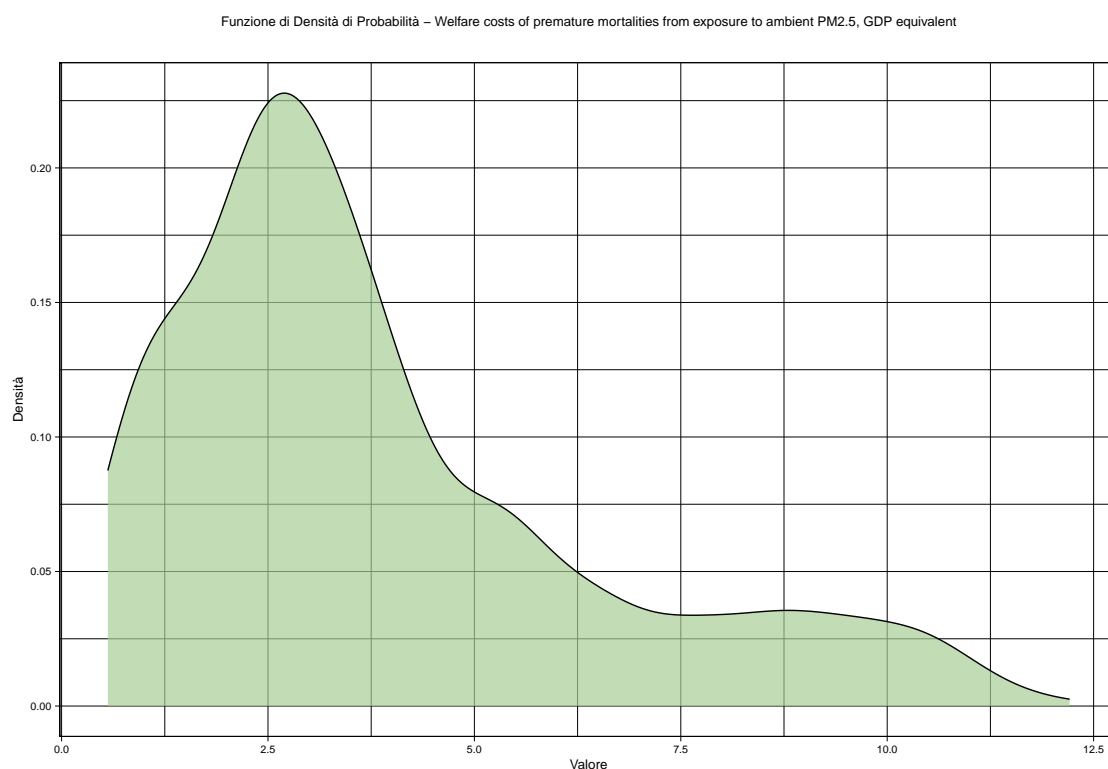


Figura 3.67: Funzione di densità

nelle condizioni ambientali, nella densità della popolazione, nelle politiche di gestione dell'inquinamento e nelle risorse disponibili per affrontare gli effetti dell'inquinamento atmosferico. La presenza di costi del benessere elevati in alcuni paesi evidenzia la necessità di politiche ambientali efficaci e di strategie di salute pubblica per ridurre l'impatto dell'inquinamento atmosferico e i suoi onerosi costi economici e sociali.

La figura 3.68 offre un'analisi dettagliata dell'andamento dei costi del benessere, espressi come percentuale del Prodotto Interno Lordo (PIL), dovuti alle mortalità precoci causate dall'esposizione a Particulate Matter 2.5 (PM2.5) nei paesi del **continente americano**. Ogni paese presenta un trend distintivo, riflettendo le diverse sfide ambientali e le risposte politiche in termini di inquinamento atmosferico.

L'**Argentina** inizia con un costo del benessere di circa il 4.3% del Prodotto Interno Lordo (PIL), seguito da una diminuzione fino a meno del 3% nel 2006. Tuttavia, dopo il 2006, i costi risalgono fino a ritornare a circa il 3% nel 2018. Questo andamento fluttuante potrebbe riflettere variazioni nelle politiche ambientali, nelle condizioni atmosferiche o nelle pratiche industriali che influenzano la qualità dell'aria.

Il **Canada** mostra una tendenza positiva con una riduzione sostanziale, passando dal 2% a quasi zero. Questo trend indica un successo significativo nel ridurre l'impatto dell'inquinamento atmosferico sulla salute pubblica e, di conseguenza, i costi economici associati.

Gli **Stati Uniti** registrano una rapida diminuzione dei costi del benessere, passando dal 4% al

1.4%. Questa marcata riduzione suggerisce miglioramenti efficaci nella qualità dell'aria e una diminuzione dell'impatto dell'inquinamento atmosferico sulla salute pubblica.

Il **Brasile** mostra un andamento relativamente stabile, con costi del benessere che oscillano intorno al 2.5%. Questo indica una situazione costante nel tempo, suggerendo che non ci sono stati cambiamenti significativi nella gestione dell'inquinamento atmosferico o nel suo impatto sulla salute pubblica.

Il **Messico** inizia con un costo del benessere del 3%, seguito da una riduzione a 2.5% nel 2007. Tuttavia, i costi risalgono successivamente, tornando al 3% nel 2018. Questo andamento suggerisce una certa volatilità nelle politiche ambientali o nelle condizioni atmosferiche, con un impatto variabile dell'inquinamento atmosferico sulla salute pubblica e sull'economia.

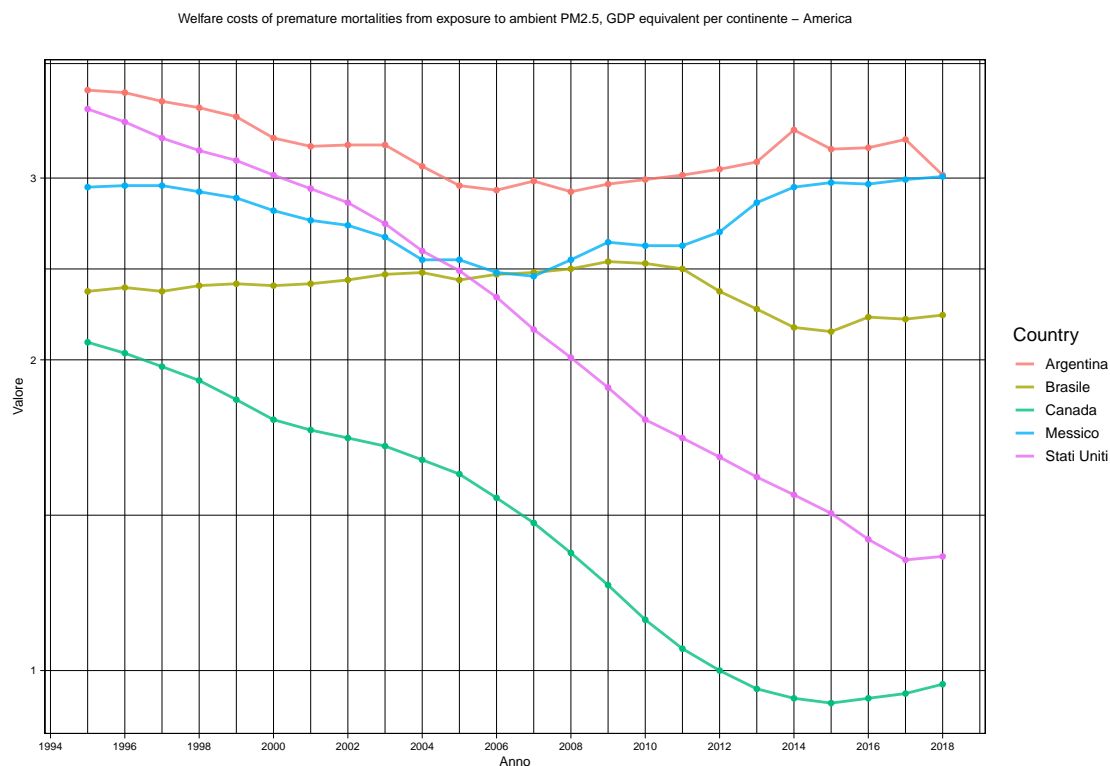


Figura 3.68: Serie Storica per il continente Americano

La figura 3.69 illustra l'andamento dei costi del benessere, espressi come percentuale del Prodotto Interno Lordo (PIL), dovuti alle mortalità precoci causate dall'esposizione a Particulate Matter 2.5 (PM2.5) nei principali **paesi asiatici**. L'andamento in ciascun paese riflette la sua unica traiettoria nella gestione dell'inquinamento atmosferico e le sue conseguenze economiche.

La **Cina** mostra un andamento inizialmente crescente, passando dal 5% al 7%, seguito da una leggera diminuzione al 6% nel 1998. Tuttavia, successivamente si assiste a una crescita marcata, con i costi che superano il 10%. Questa tendenza indica un aumento significativo dell'impatto economico dell'inquinamento atmosferico, probabilmente dovuto alla rapida industrializzazione e

urbanizzazione.

La **Russia** inizia con un alto costo del benessere del 10% e mostra una decrescita fino al 5%, nonostante alcuni periodi di incremento. Questo trend di riduzione suggerisce miglioramenti nella gestione della qualità dell'aria, sebbene con una certa variabilità nel tempo.

Il **Giappone** mantiene un livello relativamente stabile e basso, oscillando intorno al 2% e aumentando gradualmente fino al 3%. Questa stabilità riflette un impegno costante e efficace nel controllo dell'inquinamento atmosferico.

La **Corea del Sud** mostra un andamento piuttosto stabile intorno al 4%, sebbene con alcune fluttuazioni che includono sia aumenti che diminuzioni. Questo potrebbe indicare un equilibrio tra miglioramenti e sfide persistenti nella gestione dell'inquinamento atmosferico.

L'**India** presenta un rapido e marcato aumento dei costi del benessere, passando dal 4% all'8%. Questo incremento rapido potrebbe essere attribuito a un'accelerazione dell'inquinamento atmosferico in concomitanza con la rapida crescita economica e urbanizzazione.

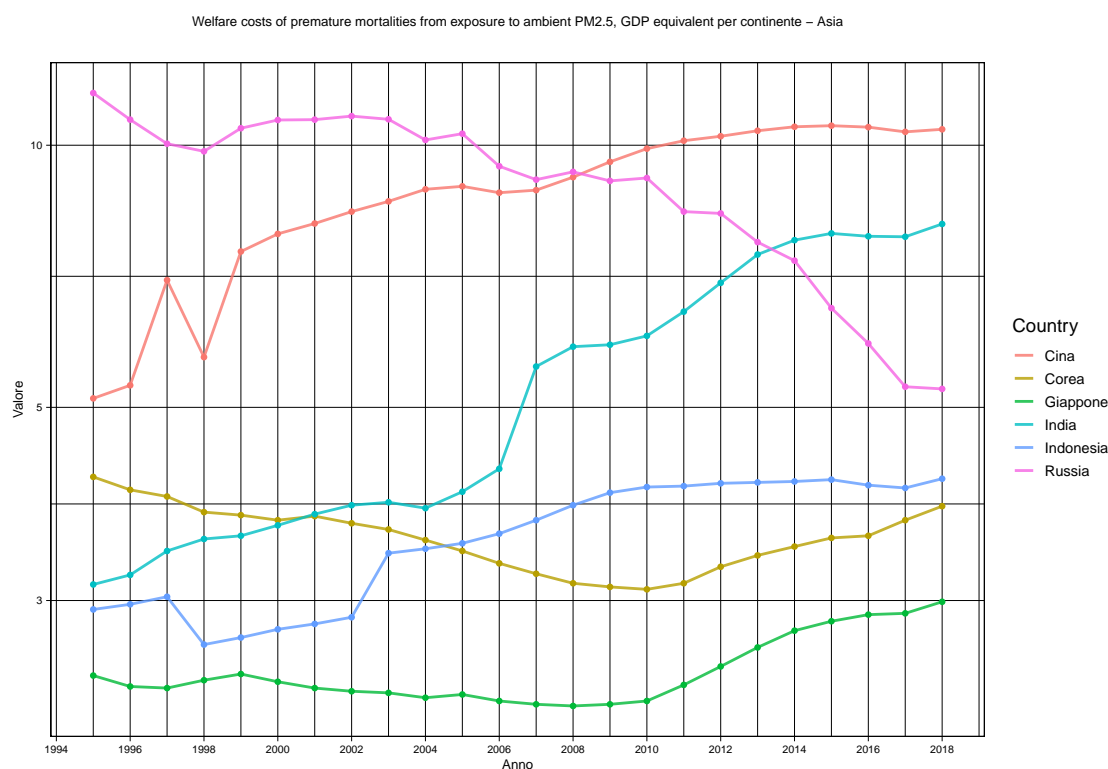


Figura 3.69: Serie storica per il continente Asiatico

La figura 3.70 fornisce un'analisi dettagliata dei costi del benessere, espressi come percentuale del Prodotto Interno Lordo (PIL), legati alle mortalità precoci dovute all'esposizione a Particulate Matter 2.5 (PM2.5) nei **paesi dell'Europa Occidentale**. Ogni paese mostra un andamento distintivo, riflettendo diverse strategie e efficacia nelle politiche di gestione della qualità dell'aria.

L'**Italia** mostra una rapida diminuzione dei costi del benessere, passando dal 7% al 4%. Tuttavia, l'andamento non è costante, con alcuni periodi di stallo. Questo potrebbe indicare fasi in cui i miglioramenti nella qualità dell'aria si sono fermati o ridotti, forse a causa di fattori economici o politici.

Il **Regno Unito** evidenzia un'impressionante riduzione dei costi, da circa il 7% al 2.5%, in un modo relativamente lineare e senza significativi stalli. Questo trend positivo suggerisce un impegno efficace e continuo nel ridurre l'inquinamento atmosferico e il suo impatto sulla salute pubblica.

La **Germania** mostra un percorso simile al Regno Unito, riducendo i costi del benessere dal 7% al 3%. Tuttavia, a differenza del Regno Unito, la Germania ha sperimentato periodi di risalita. Questi aumenti potrebbero riflettere sfide occasionali nella gestione dell'inquinamento atmosferico.

Il **Lussemburgo** presenta una notevole riduzione, passando dal 4% all'1%. Questo andamento suggerisce un successo significativo nella gestione dell'inquinamento atmosferico e una riduzione dell'impatto economico delle morti premature.

La **Spagna** mostra una riduzione moderata, passando dal 3% al 2%, ma con alcuni periodi di aumento. Questi aumenti potrebbero indicare sfide nella gestione costante dell'inquinamento atmosferico o impatti stagionali sulla qualità dell'aria.

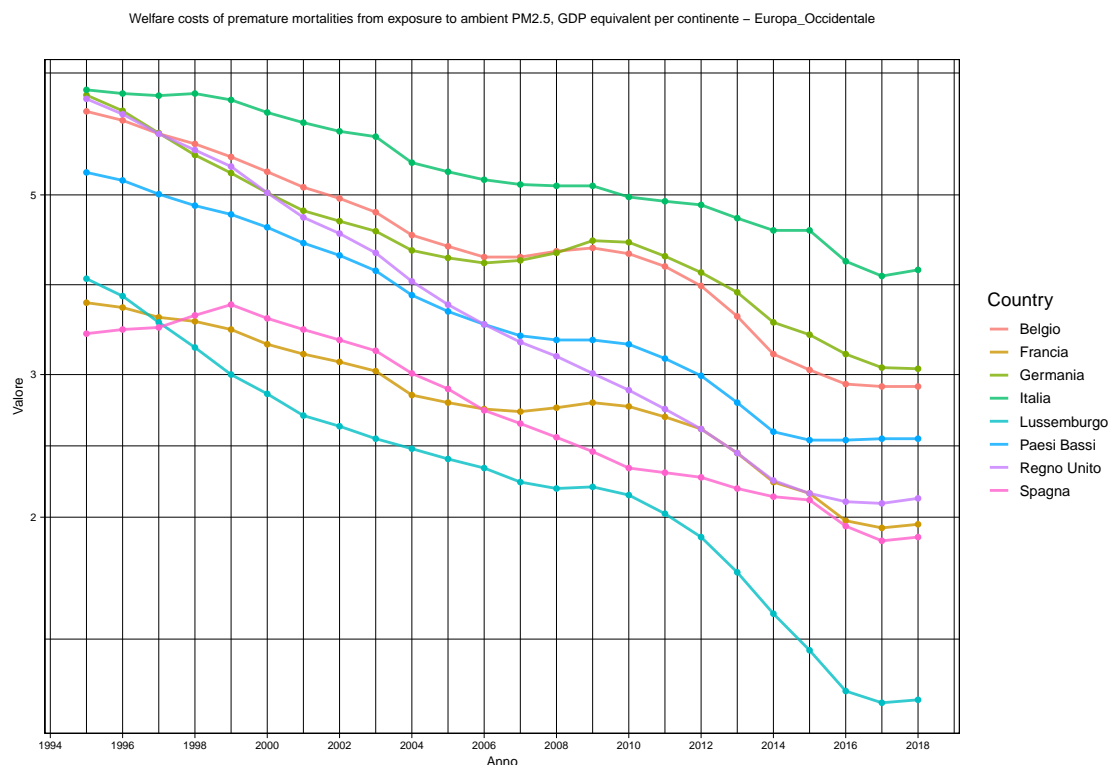


Figura 3.70: Serie storica per i paesi dell'Europa Occidentale

La figura 3.71 fornisce un'analisi dettagliata dei costi del benessere, espressi come percentuale del Prodotto Interno Lordo (PIL), legati alle mortalità precoci dovute all'esposizione a Particulate

Matter 2.5 (PM2.5) nei paesi dell'**Europa Orientale**. Questi dati offrono uno sguardo approfondito sulle sfide ambientali specifiche e sull'efficacia delle politiche in materia di qualità dell'aria in questi paesi.

L'**Ungheria** mostra un'importante riduzione dei costi, passando dal 12.5% al 7.5%, sebbene con periodi alternati di crescita e decrescita. Questo andamento fluttuante potrebbe indicare fasi di miglioramento e peggioramento nelle strategie di controllo dell'inquinamento atmosferico o potrebbe riflettere variazioni stagionali o annuali nella qualità dell'aria.

La **Polonia** segue un percorso simile all'Ungheria, riducendo i costi dal 11% al 7.5%, ma anche qui si notano periodi di aumento e diminuzione. Ciò suggerisce sfide nella gestione costante dell'inquinamento e potenziali impatti variabili sulle condizioni di salute pubblica.

La **Grecia** vede una riduzione moderata dei costi, passando dal 7.5% al 6%, ma con una tendenza altalenante. Questo potrebbe riflettere l'impatto di fattori economici e ambientali variabili sulla qualità dell'aria.

La **Danimarca** mostra una notevole diminuzione, passando rapidamente dal 6% al 2.5%. Questo trend efficace suggerisce un controllo ambientale riuscito e un significativo miglioramento nella qualità dell'aria.

Il **Portogallo** passa dal 4% al 2.5%, indicando miglioramenti progressivi nella gestione dell'inquinamento atmosferico. L'**Irlanda** registra una riduzione impressionante, passando dal 3.5% all'1.25%, il che implica un impegno efficace e continuo nella riduzione dell'esposizione a Particulate Matter 2.5 (PM2.5). Questi paesi mostrano una riduzione da 2.5% a 1.5%, con andamenti simili e sovrapponibili. Questo indica una gestione abbastanza uniforme e un impegno costante nel ridurre l'impatto dell'inquinamento atmosferico.

Mentre alcuni paesi come la **Danimarca e l'Irlanda** hanno mostrato miglioramenti rapidi e consistenti, altri come l'**Ungheria e la Polonia** hanno sperimentato andamenti più fluttuanti. Questi risultati evidenziano l'importanza di strategie ambientali efficaci e adattabili per affrontare l'inquinamento atmosferico e ridurre l'impatto economico e sulla salute.

La figura 3.72 offre una panoramica del trend dei costi del benessere, espressi come percentuale del Prodotto Interno Lordo (PIL), associati alle mortalità precoci dovute all'esposizione a Particulate Matter 2.5 (PM2.5) nei paesi di **Oceania e Africa**, in particolare Australia, Nuova Zelanda e Sud Africa.

Sia l'**Australia che la Nuova Zelanda** mostrano una riduzione significativa dei costi del benessere, passando da circa il 1.5% a quasi lo 0%. Questo andamento quasi parallelo tra i due paesi suggerisce un efficace controllo dell'inquinamento atmosferico e un'impressionante gestione della qualità dell'aria. Il trend in calo costante indica che sono state prese misure efficaci per ridurre l'impatto sanitario dell'inquinamento atmosferico, probabilmente attraverso politiche ambientali rigorose e sostenibili.

Il **Sud Africa**, al contrario, mostra un incremento dei costi del benessere, da circa il 4% al 5%. Tuttavia, l'andamento non è lineare e si caratterizza per periodi alternati di crescita e decrescita. Questo potrebbe riflettere sfide persistenti nella gestione dell'inquinamento atmosferico, forse

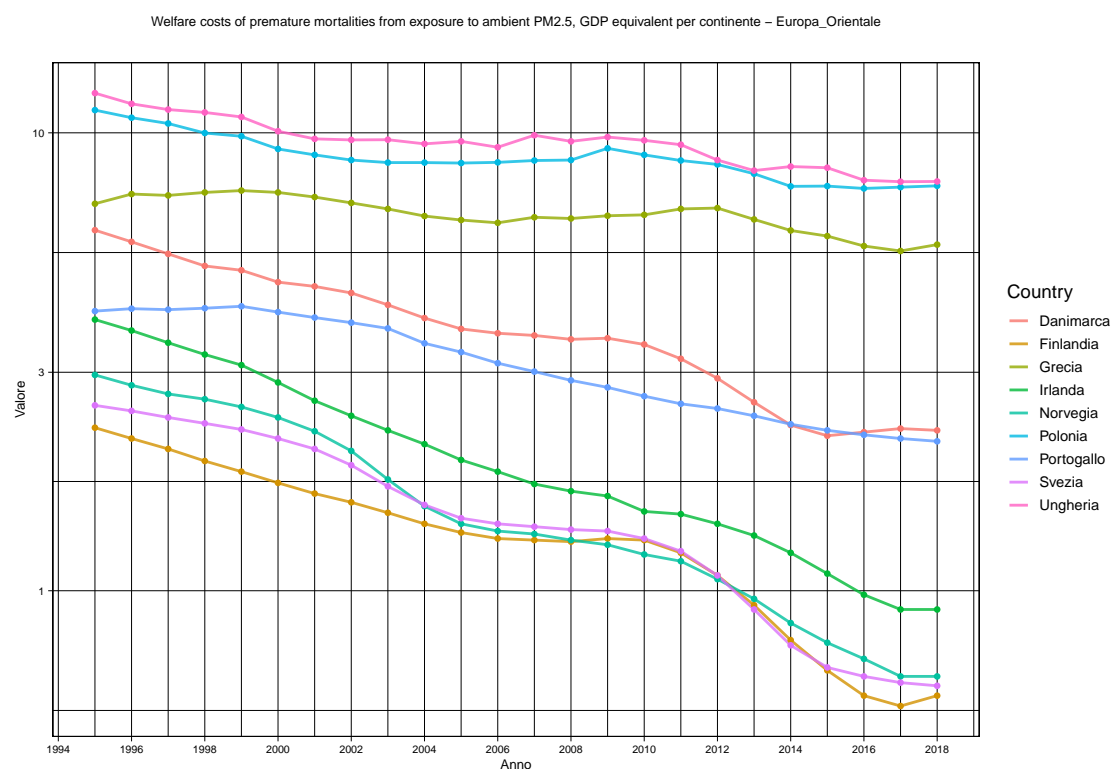


Figura 3.71: Serie storica per i paesi dell'Europa Orientale

dovute a variazioni nelle pratiche industriali, nelle politiche ambientali o nelle condizioni meteorologiche. L'aumento complessivo dei costi del benessere indica che l'impatto dell'inquinamento atmosferico sulla salute pubblica e sull'economia rimane una questione preoccupante in questa regione.

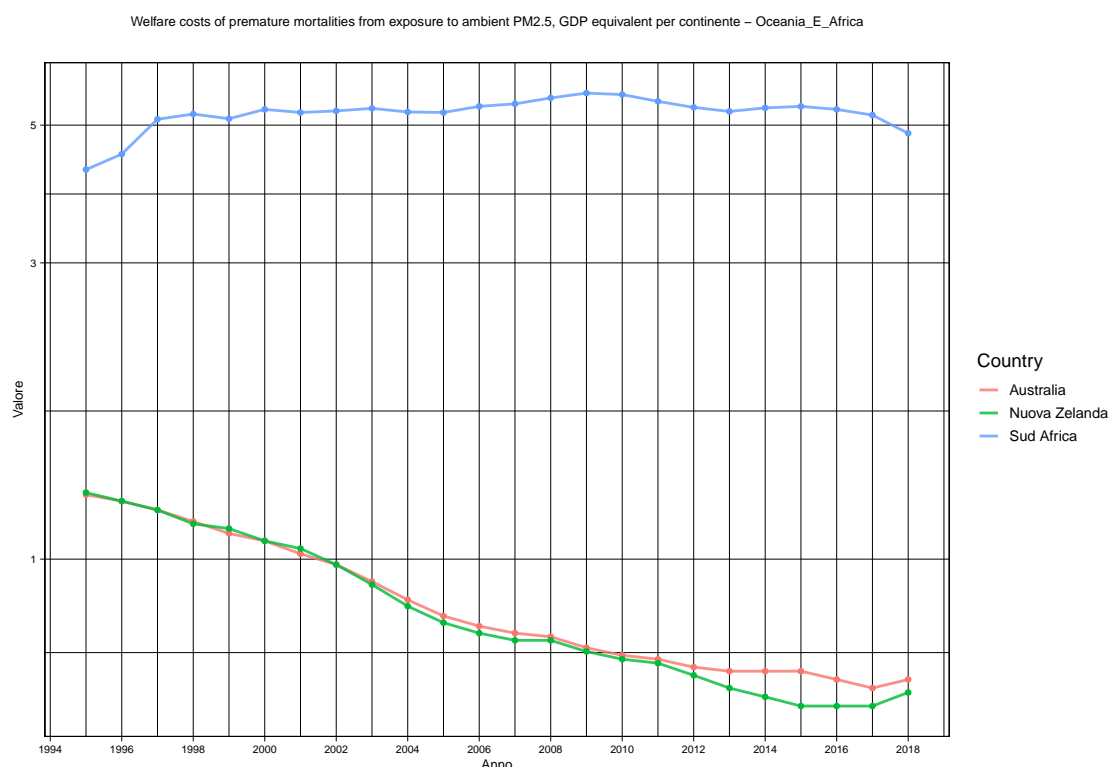


Figura 3.72: Serie storica per i paesi dell' Africa e dell'Oceania

3.2.10 Area Marina Protetta % superficie terrestre

L'area marina protetta è espressa come percentuale della zona economica esclusiva totale. Le aree protette includono tutte le aree con una categoria di gestione assegnata dall'Unione Internazionale per la Conservazione della Natura (IUCN). Le aree con categorie di gestione sovrapposte sono conteggiate una sola volta per evitare il doppio conteggio. Tuttavia, alcuni paesi segnalano alcune aree protette come punti (senza dettagli sulla posizione georeferenziata), in questi casi le sovrapposizioni non possono essere identificate o risolte. La superficie totale delle aree protette presentata qui include queste aree segnalate come punti. Zona Economica Esclusiva (ZEE) di un paese comprende, in base alla Convenzione delle Nazioni Unite del 1982 sul Diritto del Mare, 200 miglia marine dalla costa o il punto mediano tra le coste quando le ZEE di diversi paesi si sovrappongono. Alcune aree marine sono condivise tra paesi in regimi congiunti o ci sono contestazioni non risolte per sovrapposizione di rivendicazioni su un'area. In questi casi, l'intera area rilevante (e le aree protette al suo interno) sono attribuite a tutti i paesi interessati. Per ulteriori dettagli sulla metodologia, si prega di consultare gli indicatori sulle aree protette terrestri e marine: metodologia e risultati per i paesi dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE) e del G-20. I dati sulle aree protette sono ottenuti dal dataset delle Aree Protette nel Database sull'Ambiente dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE).

In **media**, le aree marine protette rappresentano il 7.54% della superficie terrestre. È una misura centrale che riflette la tendenza generale. La **mediana** (2.27) essendo significativamente inferiore alla media, indica che più della metà dei paesi ha una percentuale di area marina protetta inferiore al 2.27% della loro superficie terrestre. Questo suggerisce una **distribuzione asimmetrica**, confermato dal valore di *swness* di (1.932312), Un valore maggiore di 1 indica una distribuzione

Area Marina Protetta % superficie terrestre	
Media	7.54
Mediana	2.27
Simmetria	1.932312
Curtosi	3.164073
Frequenza assoluta	744
Massimo	45.6
Minimo	0
Deviazione standard	10.77089
coefficiente di variazione	1.429433
Varianza	116.0122
Q1	0.677
Q3	10.1

Tabella 3.10: Statiche descrittive per la variabile: Area Marina Protetta % superficie terrestre

asimmetrica positiva, con una coda che si estende verso i valori più alti. Il valore di **curtosi** suggerisce che la distribuzione ha code più pesanti rispetto a una distribuzione normale, indicando la presenza di valori estremi.

Il **massimo** (45.6) e il **minimo** (0) indicano l'intervallo dei dati, con il valore massimo di area marina protetta come percentuale della superficie terrestre raggiungendo il 45.6%, mentre il minimo è 0%. Una **deviazione standard** elevata suggerisce una grande variabilità nei dati, confermando la presenza di valori estremi. Il **Coefficiente di Variazione** (1.429433) indica una variabilità molto elevata in proporzione alla media.

Il Primo Quartile (Q1) indica che il 25% dei dati ha una percentuale di area marina protetta inferiore a 0.677%, mentre il Terzo Quartile (Q3) suggerisce che il 75% dei dati è sotto il 10.1%, sottolineando ulteriormente la distribuzione asimmetrica dei dati.

Il **boxplot** 3.73 mostra una mediana (2.27%) significativamente inferiore alla media (7.54%), indicando una distribuzione asimmetrica. L'**asimmetria positiva** (1.932312) è confermata dal fatto che la maggior parte dei dati è concentrata verso il valore minimo (0%), con alcune osservazioni estreme che si spingono fino al 45.6%. La distanza tra il Primo Quartile (Q1), (0.677%) e il Terzo Quartile (Q3) (10.1%) suggerisce una variabilità considerevole nella percentuale di area marina protetta tra i diversi paesi.

Il grafico 3.74 mostra l'**istogramma** che conferma ulteriormente l'asimmetria positiva, con una concentrazione di valori vicino allo zero e poche osservazioni ad alto valore. La distribuzione non è uniforme, e ci sono poche nazioni con una percentuale molto alta di aree marine protette rispetto alla superficie terrestre.

Il **grafico della densità** 3.75 mostra una curva che enfatizza l'asimmetria e la curtosi (3.164073) della distribuzione. L'alta curtosi indica una distribuzione con una coda più pesante rispetto a una distribuzione normale, suggerendo la presenza di valori estremi (come il valore massimo di 45.6%)

Nelle figure 3.76 , 3.77, 3.78, 3.79, 3.80 sono raffigurate la serie storica per la variabili per i vari paesi considerati nel periodo in esame 1995/2018. Sull'asse delle ascisse sono rappresentati gli anni mentre, sull'asse delle ordinate sono rappresentati i valori delle percentuale della area mariane

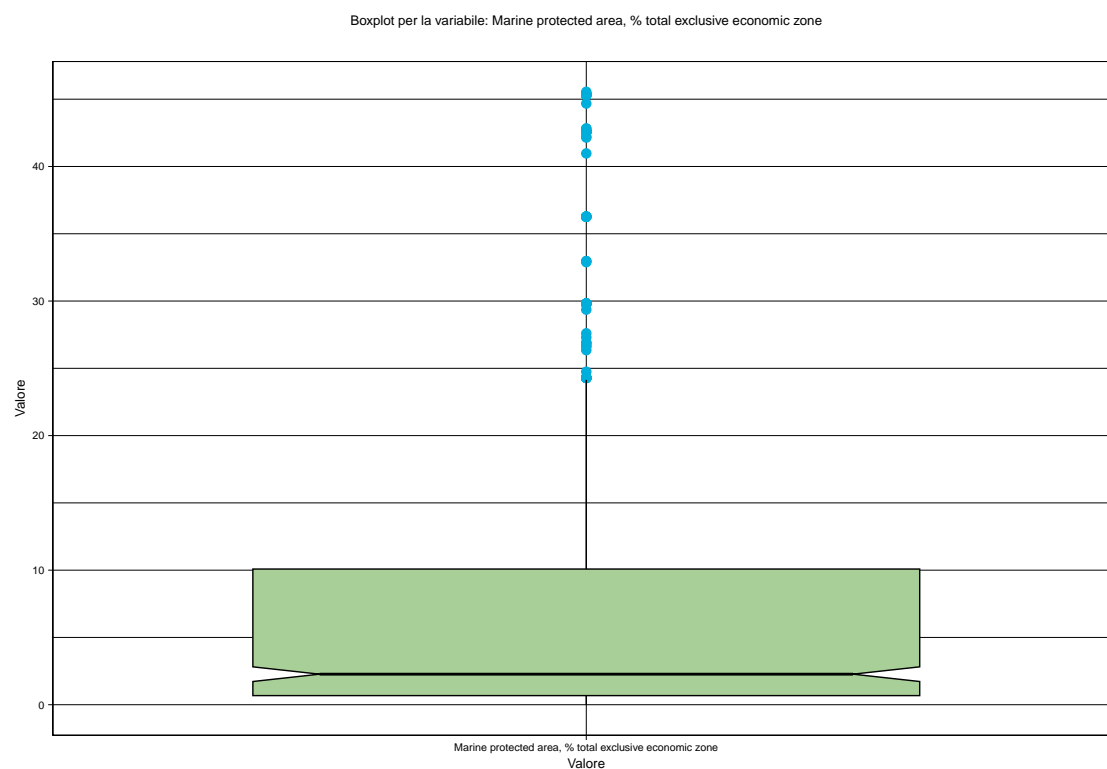


Figura 3.73: Boxplot per la variabile:Area Marina Protetta % superficie terrestre

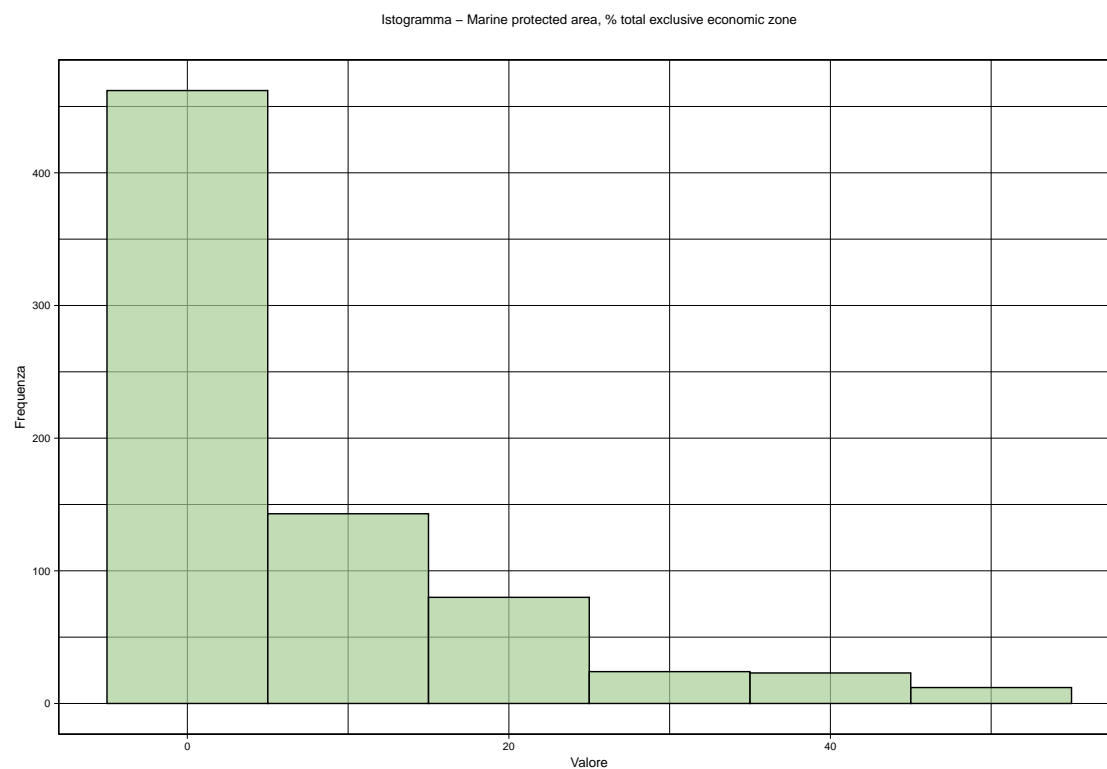


Figura 3.74: Istogramma

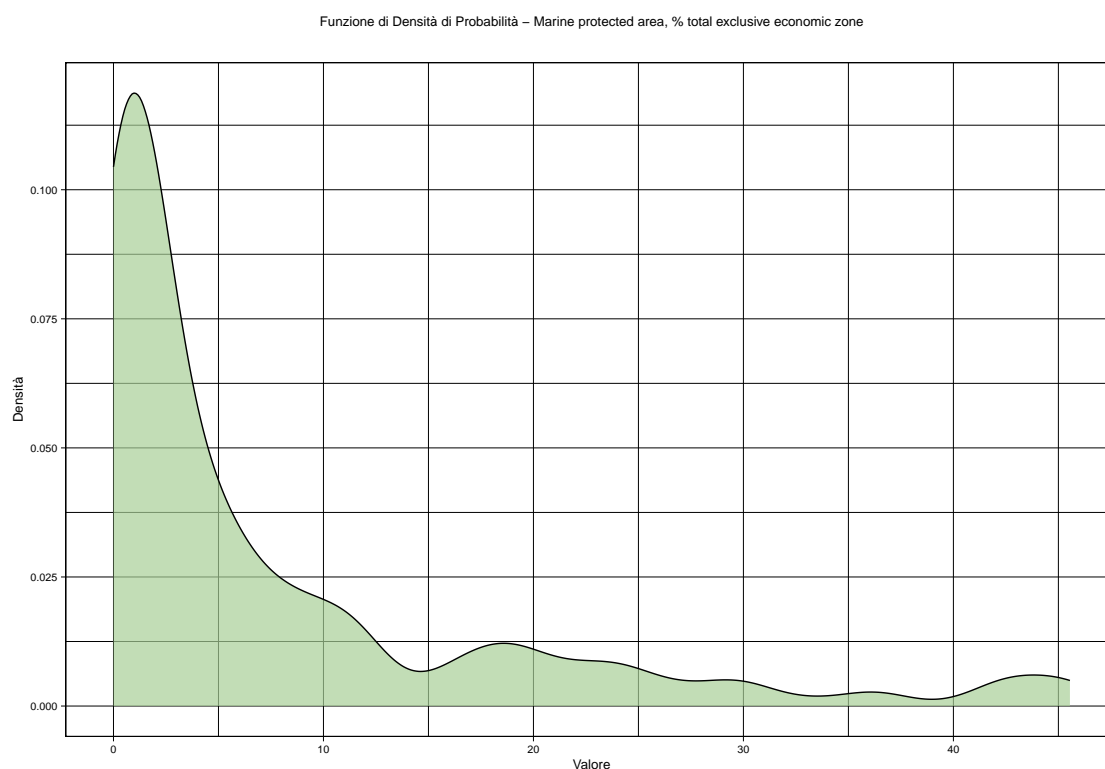


Figura 3.75: Istogramma

potrete. All'interno del grafico ogni linea rappresenta una serie storica di un paese.

La figura 3.76 illustra in modo efficace l'evoluzione della percentuale delle aree marine protette nei paesi del **continente Americano**. Il grafico mostra chiaramente che, fino al 2004, la protezione delle aree marine non era una priorità evidente per questi paesi, come si può dedurre dalla quota vicina allo zero. Questo suggerisce che, in quel periodo, le iniziative per la conservazione marina erano in uno stato embrionale o che le aree protette esistenti non erano sufficientemente riconosciute o formalmente designate.

A partire dal 2004, tuttavia, si osserva un significativo cambiamento, con una rapida e consistente crescita nella percentuale delle aree marine protette. Questo cambiamento marcato potrebbe essere il risultato di un crescente riconoscimento dell'importanza della conservazione marina, sia a livello nazionale che internazionale. È possibile che iniziative globali o regionali, come accordi ambientali o conferenze internazionali, abbiano stimolato questa svolta, incentivando i paesi americani a dedicare maggiori risorse e attenzione alla protezione degli ecosistemi marini.

Questo aumento della percentuale delle aree marine protette rappresenta un passo positivo verso la conservazione della biodiversità marina e la protezione degli habitat critici. È indicativo di un cambiamento nell'atteggiamento e nelle politiche ambientali, che pone ora maggiore enfasi sulla necessità di salvaguardare gli oceani per la loro importanza ecologica, economica e scientifica.

Esaminando la figura 3.77, che illustra le tendenze delle aree marine protette in diversi **paesi asiatici**, si nota una gamma di approcci e risultati. Questi riflettono non solo le politiche ambientali

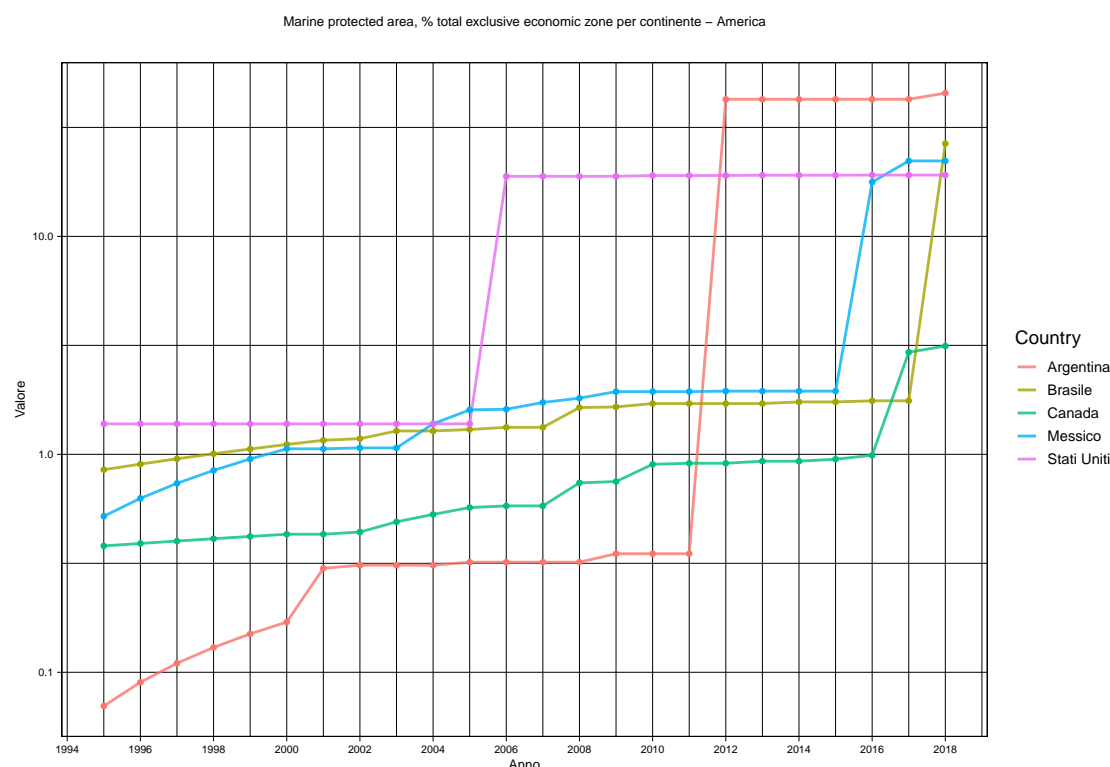


Figura 3.76: Serie storica per il continente Americano

di ciascun paese, ma anche i loro contesti economici, geografici e sociali unici.

Il **Giappone** ha un leggero aumento da 7.5% a circa 8% nelle aree marine protette mostra un impegno continuo, seppur moderato, nella conservazione marina. Questo incremento, è indicativo di una politica ambientale coerente, forse influenzata dalla ricca biodiversità marina del Giappone e dalla sua storica connessione con il mare.

Passando all'**India**, la situazione appare notevolmente diversa. La costanza nel valore di 0% potrebbe suggerire che la conservazione marina non è stata una priorità significativa. Questo potrebbe essere dovuto a una serie di fattori, come limitazioni di risorse, altre priorità di sviluppo, o sfide nella gestione e implementazione di politiche ambientali.

In **Corea**, la stabilità attorno all'1.25% può essere interpretata come un impegno costante, ma limitato, verso la conservazione marina. Questo valore stabile potrebbe riflettere una situazione in cui sono state create aree marine protette, ma senza significative espansioni o nuove iniziative nel periodo considerato.

La situazione in **Cina**, con valori intorno allo 0.5%, suggerisce un impegno minimo nella conservazione marina. Ciò potrebbe essere attribuito a una varietà di fattori, inclusa la rapida industrializzazione e crescita economica del paese, che potrebbero aver avuto la precedenza sulle iniziative di conservazione marina.

L'**Indonesia** si distingue in modo significativo. Con la crescita più marcata, raggiungendo circa il

3%, dimostra un impegno crescente verso la protezione dei suoi ecosistemi marini. Considerando l'incredibile biodiversità marina dell'Indonesia e il suo ruolo cruciale nell'ecosistema marino globale, questo aumento è sia importante che necessario.

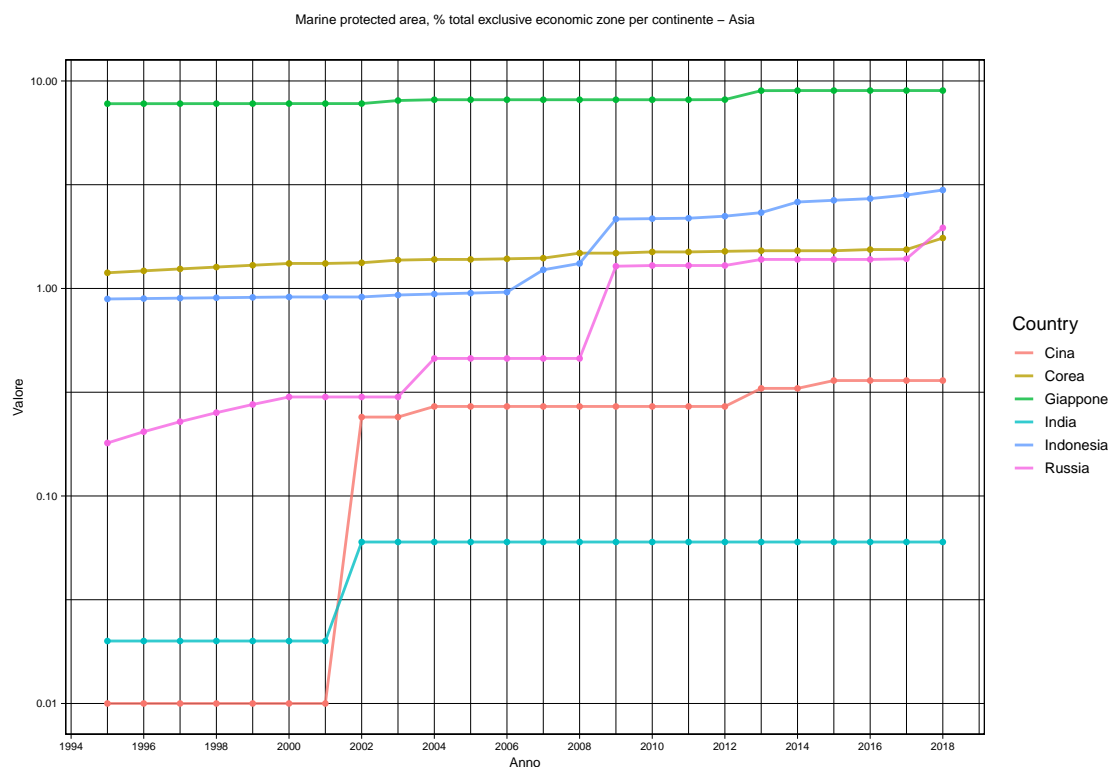


Figura 3.77: Serie storica per il continente Asiatico

La figura 3.78, che mostra le serie storiche delle aree marine protette in **Europa Occidentale**, offre un quadro interessante delle tendenze ambientali in questa regione.

Iniziamo con il **Lussemburgo**. Il valore di 0% per le aree marine protette è del tutto prevedibile, dato che il Lussemburgo è un paese senza sbocchi sul mare. Per quanto riguarda gli altri paesi dell'**Europa Occidentale**, si nota un trend generale di aumento costante e notevole nelle aree marine protette. In Europa, c'è stata una crescente consapevolezza riguardo alle questioni ambientali, inclusa la necessità di proteggere gli ecosistemi marini. Questo ha probabilmente influenzato le politiche pubbliche e le azioni di conservazione.

La figura 3.79, che mostra le serie storiche delle aree marine protette in **Europa Orientale**, offre un quadro interessante delle tendenze ambientali in questa regione.

Iniziamo con l'**Ungheria**. Il valore di 0% per le aree marine protette è del tutto prevedibile, dato che il paese è un paese senza sbocchi sul mare. Per quanto riguarda gli altri paesi dell'Europa Orientale, è interessante notare un trend generale di aumento nelle aree marine protette, sebbene in misura minore rispetto all'Europa Occidentale.

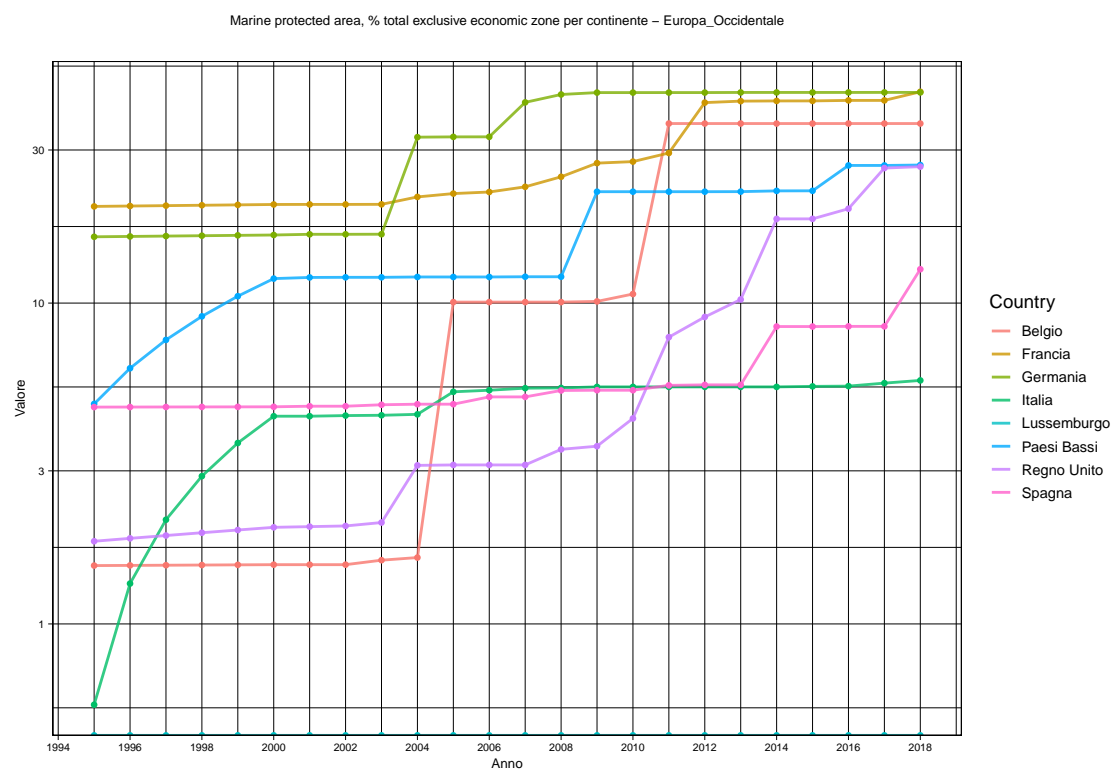


Figura 3.78: Serie storica per i paesi dell'Europa Occidentale

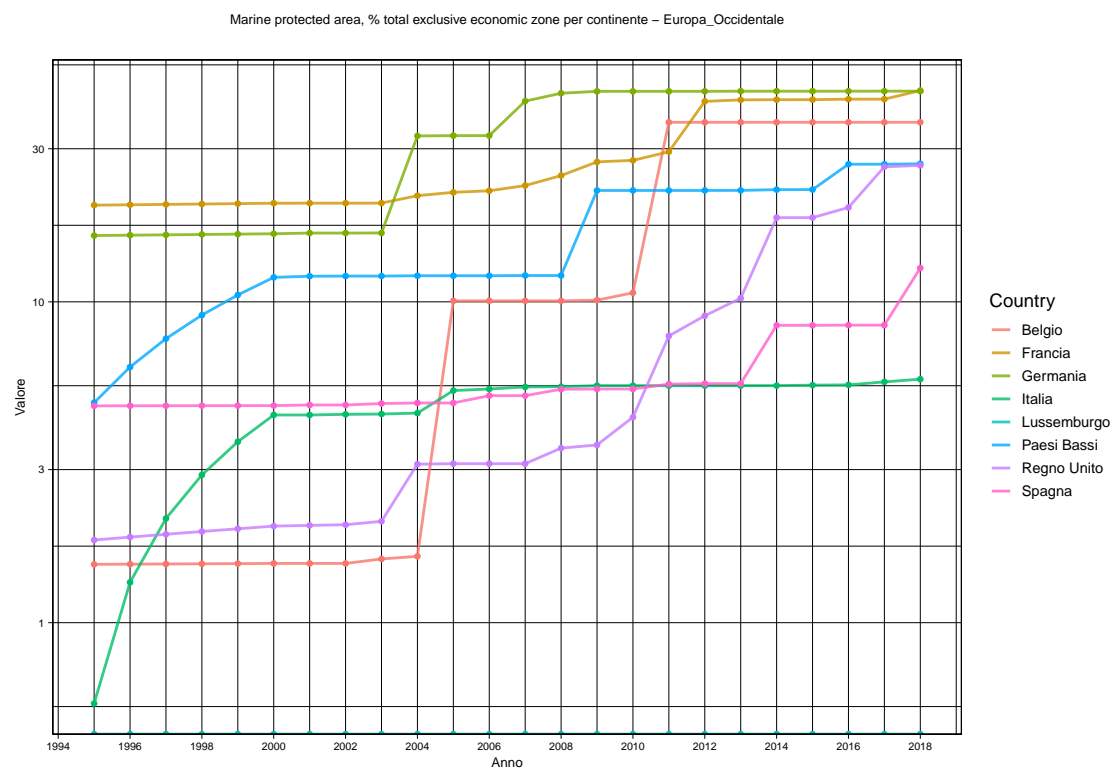


Figura 3.79: Serie storica per i paesi dell'Europa Orientale

La figura 3.80 ci offre una panoramica importante sui progressi fatti dai paesi di **Africa e Oceania** nella conservazione delle aree marine. L'analisi delle tendenze in questi paesi rivela degli sviluppi significativi e rapidi nel campo della protezione degli ambienti marini.

L'**Australia**, mostra un notevole progresso passando dal 5% al 40% di aree marine protette. Questo incremento è particolarmente impressionante e riflette una forte dedizione alla conservazione degli ecosistemi marini. L'Australia, con la sua vasta linea costiera e unico patrimonio marino, ha chiaramente riconosciuto l'importanza di proteggere le sue risorse naturali. Guardando al Sud Africa e alla Nuova Zelanda, entrambi i paesi partono da una baseline del 0% e raggiungono rispettivamente il 10% e il 30%. Questi aumenti, sebbene meno drastici rispetto all'Australia, sono tuttavia significativi.

Il **Sud Africa e la Nuova Zelanda** hanno entrambi coste estese e diversificate, e il loro crescente impegno nella protezione delle aree marine potrebbe essere influenzato da una consapevolezza crescente delle minacce ambientali, come il cambiamento climatico e la sovrapesca, oltre che da un impegno generale verso la sostenibilità ambientale. La rapidità di questi aumenti è degna di nota. Potrebbe indicare un cambiamento significativo nelle politiche ambientali, forse stimolato da eventi catalizzatori come conferenze internazionali sulla biodiversità, report scientifici allarmanti sull'ambiente marino, o campagne pubbliche di sensibilizzazione. Questo tipo di crescita rapida spesso si verifica in risposta a pressioni ambientali acute o a un cambiamento nella percezione pubblica e politica dell'importanza della conservazione marina.

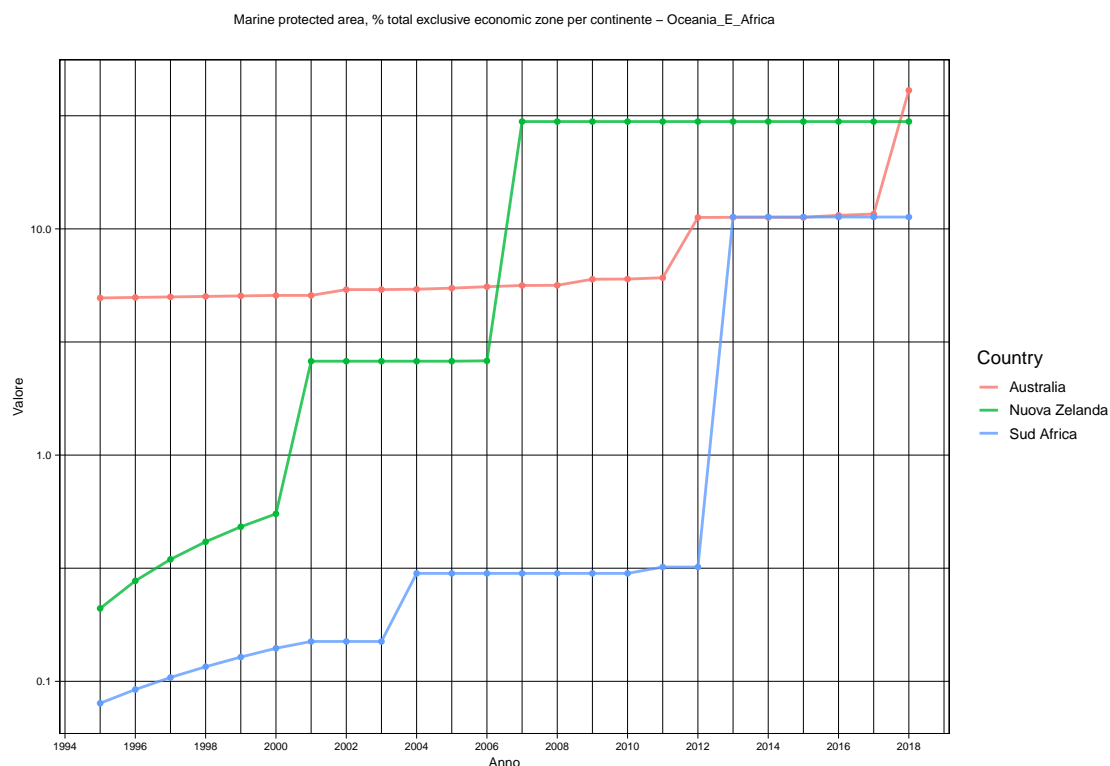


Figura 3.80: Serie storica per i paesi dell'Africa e dell'Oceania

3.2.11 Area Terrestre Protetta % superficie terrestre

L'area terrestre protetta è espressa come percentuale della superficie terrestre totale. Le aree protette includono tutte le aree con una categoria di gestione assegnata dall'Unione Internazionale per la Conservazione della Natura (IUCN). Le aree con categorie di gestione sovrapposte sono conteggiate una sola volta per evitare il doppio conteggio. Tuttavia, alcuni paesi segnalano alcune aree protette come punti (senza dettagli sulla posizione georeferenziata), in questi casi le sovrapposizioni non possono essere identificate o risolte. La superficie totale delle aree protette presentata qui include queste aree segnalate come punti. Per ulteriori dettagli sulla metodologia, si prega di consultare gli Indicatori sulle Aree Protette Terrestri e Marine: Metodologia e Risultati per i Paesi dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE) e del G20. I dati sulle aree protette sono ottenuti dal dataset delle Aree Protette nel Database sull'Ambiente dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE).

Area Terrestre Protetta % superficie terrestre	
Media	16.0
Mediana	13.6
Simmetria	0.9400385
Curtosi	0.6231474
Frequenza assoluta	744
Massimo	45.6
Minimo	0
Deviazione standard	10.02717
Coefficiente di variazione	0.6282943
Varianza	100.5442
Q1	9.51
Q3	21.8

Tabella 3.11: Statiche descrittive per la variabile: Area Terrestre Protetta % superficie terrestre

La **media** (16.0%) indica che in media, circa il 16% della superficie terrestre nei paesi considerati è designata come area protetta. Questo è un valore relativamente alto, che suggerisce un impegno significativo nella conservazione terrestre. La **mediana** (13.6%) più bassa della media, indica che più della metà dei paesi ha una percentuale di area terrestre protetta inferiore al 13.6%. La **simmetria** (0.9400385) un valore leggermente inferiore a 1 suggerisce una distribuzione moderatamente **asimmetrica positiva**, indicando che alcuni paesi hanno valori molto più alti rispetto alla media. La **curtosi** (0.6231474) ha un valore relativamente basso indica che la distribuzione non presenta code particolarmente pesanti. La **deviazione standard** (10.02717): Una discreta variabilità, indicando differenze significative tra i paesi nella percentuale di aree terrestri protette.

Il **boxplot** 3.81 mostra una distribuzione relativamente equilibrata tra i valori bassi e alti, con una **mediana** (13.6%) leggermente inferiore alla **media** (16.0%). L'**asimmetria leggermente positiva** (0.9400385) suggerisce che ci sono più paesi con una percentuale di aree protette superiore alla mediana rispetto a quelli al di sotto. I valori di Q1 (9.51%) e Q3 (21.8%) indicano che la metà dei paesi ha una percentuale di aree protette compresa tra questi due valori, suggerendo una varietà di approcci alla conservazione terrestre.

La **deviazione standard** (10.02717) e la **varianza** (100.5442) mostrano una variabilità significativa tra i paesi, ma non eccessiva. Il **coefficiente di variazione** (0.6282943) suggerisce che la

variabilità relativa è moderata rispetto alla media. La presenza di paesi con un valore **massimo** del 45.6% e un **minimo** di 0% mostra un ampio range nella percentuale di aree terrestri protette.

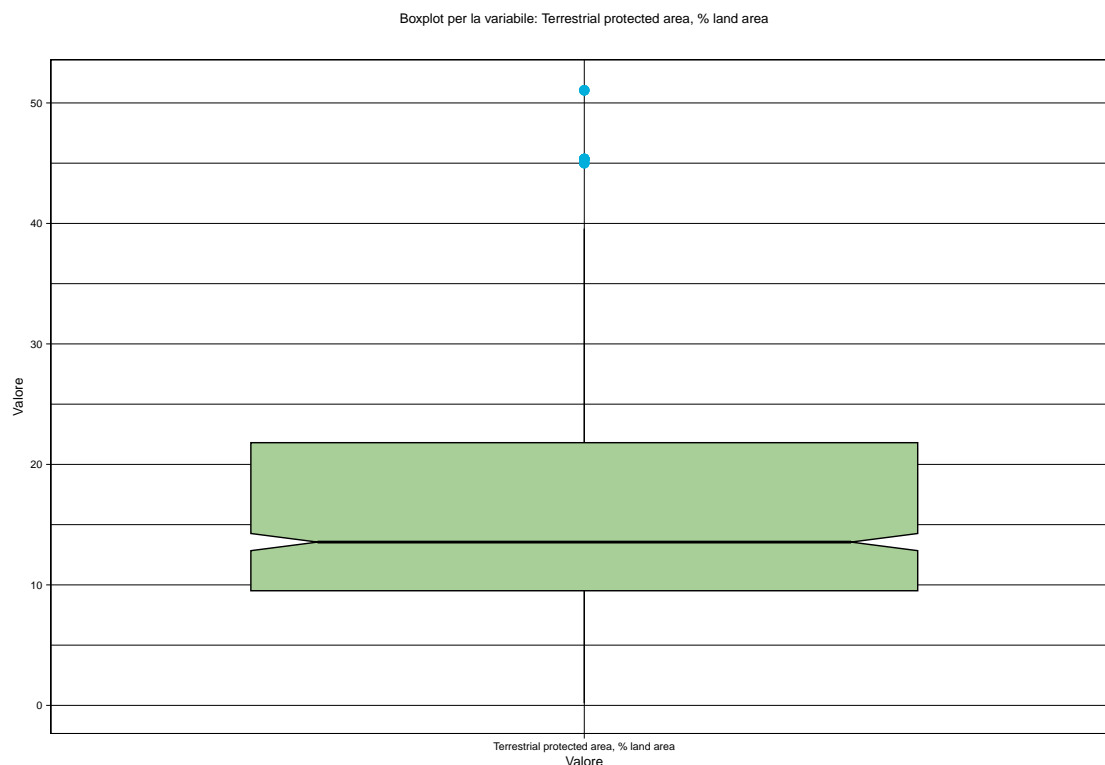


Figura 3.81: Boxplot per la variabile: Area Terrestre Protetta % superficie terrestre

L'**istogramma** conferma la distribuzione con una leggera asimmetria positiva. La maggior parte dei paesi ha una percentuale di aree terrestri protette tra 0% e 20%, con un numero minore di paesi che raggiungono percentuali molto elevate.

Il **grafico della densità** 3.83 mostra la distribuzione dei dati con una curva di densità che riflette l'asimmetria e la curtosi (0.6231474) della distribuzione. La curtosi relativamente bassa indica che la distribuzione non ha code particolarmente pesanti e si avvicina a una distribuzione normale.

La figura 3.84 offre un'analisi dettagliata della serie storica relativa alle aree terrestri protette nel **continente americano**, mettendo in luce differenze significative nell'approccio alla conservazione tra vari paesi.

Il **Brasile** mostra un notevole aumento della percentuale delle aree terrestri protette, partendo da circa il 10% e raggiungendo rapidamente il 30% entro il 2010. Questo rapido incremento potrebbe riflettere un impegno deciso del paese nella conservazione della biodiversità e nella protezione delle sue vaste risorse naturali, in particolare le foreste tropicali. Tuttavia, il plateau raggiunto dopo il 2010 suggerisce che questo impegno si è stabilizzato, forse a causa di cambiamenti nelle politiche ambientali o di altre priorità nazionali.

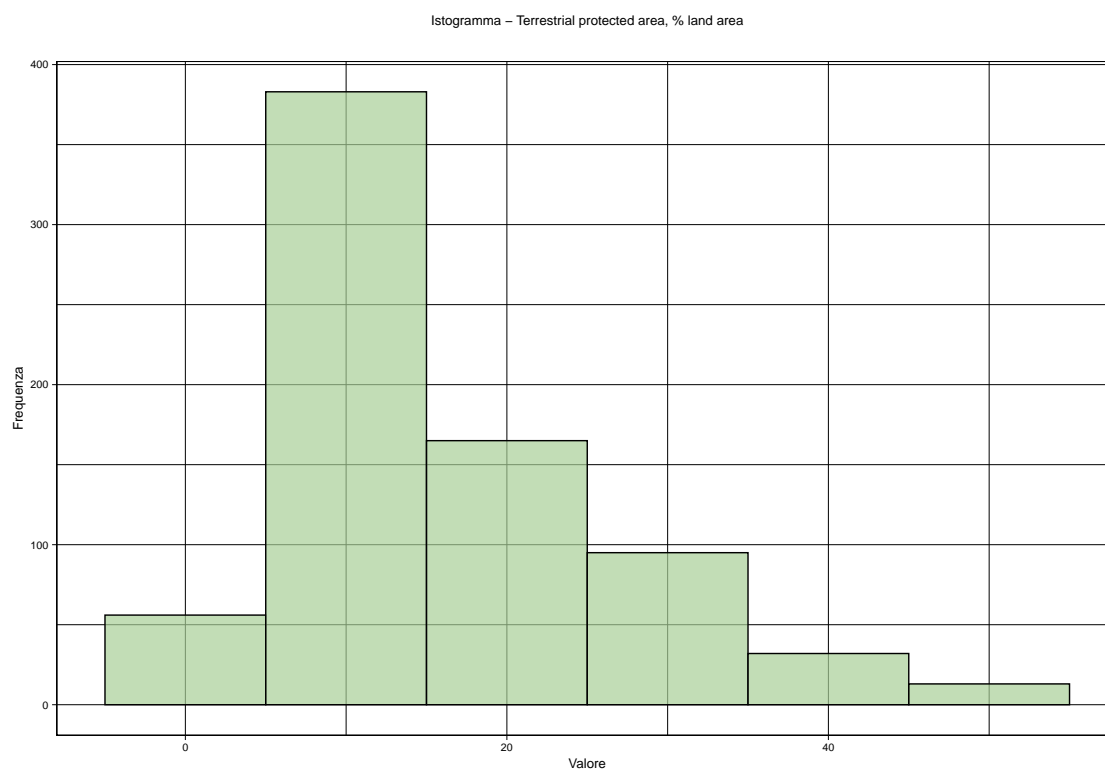


Figura 3.82: Istogramma

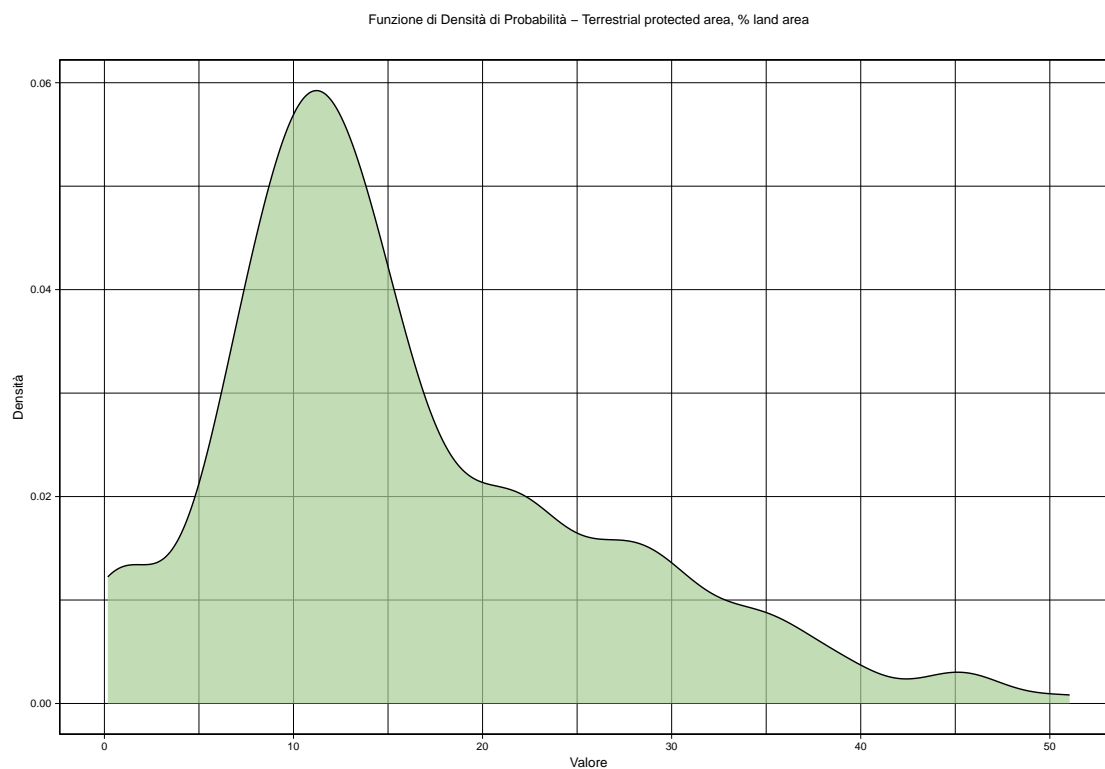


Figura 3.83: Funzione di densità

Gli **Stati Uniti** mostrano un incremento graduale e costante della percentuale delle aree protette, passando dal 10% all'11%. Questo andamento lento potrebbe indicare un approccio più graduale alla designazione di nuove aree protette.

Il **Messico** parte da una percentuale inferiore al 5% e raggiunge circa il 15%, con un aumento relativamente rapido seguito da un periodo di stallo. Questo potrebbe suggerire una fase iniziale di forte impegno per la conservazione, seguita da una fase in cui altri fattori hanno rallentato o impedito ulteriori incrementi.

Canada e Argentina mostrano andamenti simili, con percentuali quasi costanti. Questa costanza potrebbe indicare un approccio più stabile e forse più cauto alla conservazione terrestre, con minori variazioni nel corso del tempo.

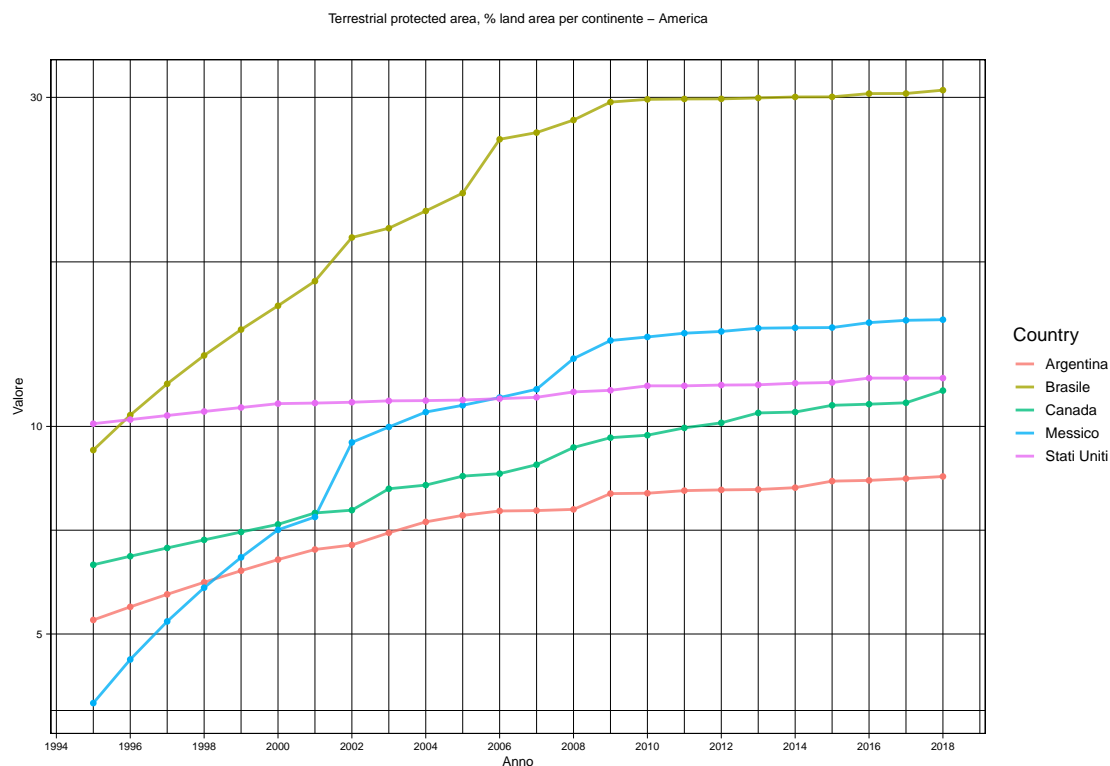


Figura 3.84: Serie storica per il continente Americano

La figura 3.85 mette in evidenza le tendenze nel settore della conservazione terrestre in alcuni dei principali **paesi asiatici**. Questi dati rivelano approcci diversi e dinamiche uniche relative alla gestione e alla protezione delle aree naturali.

Il **Giappone** mostra un incremento graduale della percentuale di aree terrestri protette, passando dal 18% al 21%. Questa crescita lenta ma costante riflette probabilmente un impegno continuo nella conservazione ambientale, in linea con l'approccio giapponese tradizionalmente prudente e sistematico nella gestione delle risorse naturali.

In **Russia**, si osserva un incremento più significativo, da una percentuale iniziale inferiore al 5% a circa l'11%. Questo andamento suggerisce un aumento dell'attenzione e dell'impegno nella conservazione terrestre nel corso degli anni. La Russia, con le sue vaste aree naturali e la ricchezza di biodiversità, ha il potenziale per influenzare significativamente la conservazione a livello globale.

L'**Indonesia** mostra un aumento da circa l'11% al 12.5%, raggiungendo questo picco nel 2011, per poi stabilizzarsi. Questo potrebbe indicare un periodo di forte impegno nella designazione di aree protette, seguito da un periodo di consolidamento o di cambiamenti nelle priorità politiche.

India e Cina presentano andamenti simili, con valori relativamente bassi, vicini allo 0%. Questo suggerisce che la conservazione terrestre potrebbe non essere stata una priorità alta per questi paesi, almeno fino a tempi recenti. Questa situazione potrebbe essere dovuta a una varietà di fattori, come lo sviluppo economico rapido, le pressioni demografiche e la necessità di bilanciare la conservazione con altri obiettivi di sviluppo.

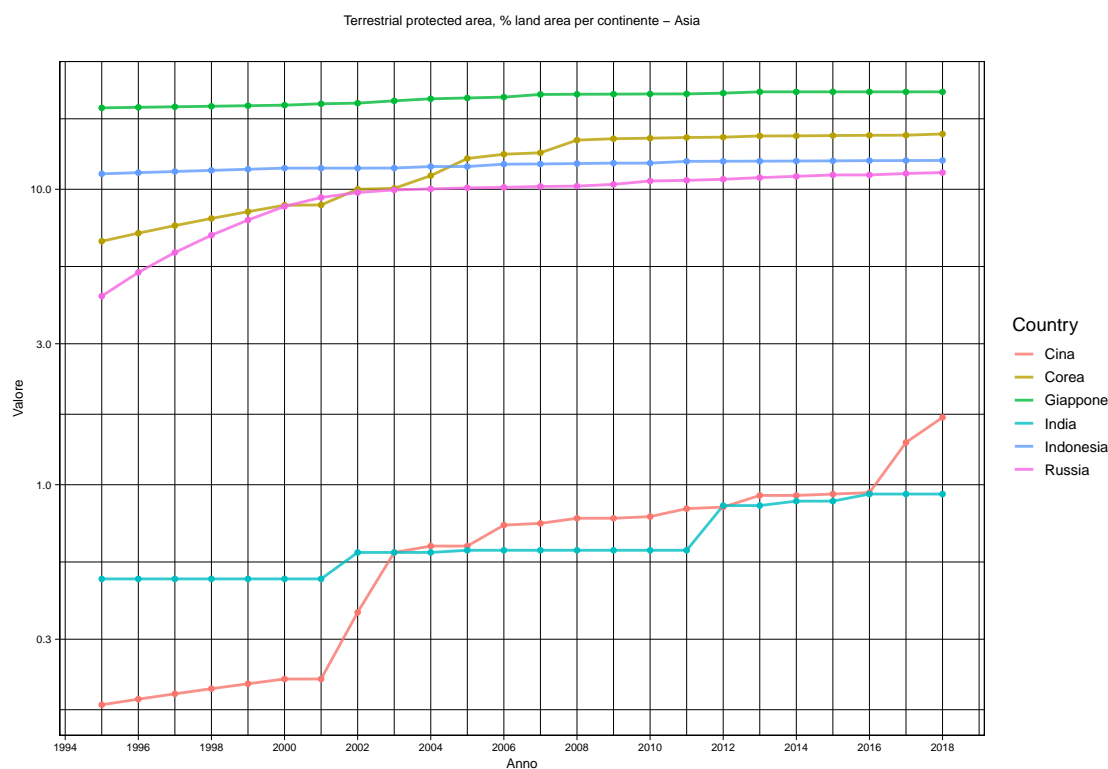


Figura 3.85: Serie storica per il continente Asiatico

La figura 3.86 rivela come diversi paesi dell'Europa Occidentale abbiano evoluto le loro politiche e pratiche riguardanti le aree terrestri protette nel corso degli anni. Il **Lussemburgo** mostra un incremento notevole, raddoppiando quasi la percentuale di aree terrestri protette, da 25% a 50%. L'andamento è caratterizzato da periodi di stallo, seguiti da fasi di crescita.

Il **Regno Unito** ha una crescita più graduale, passando dal 24% al 28%. Questo incremento lento ma costante suggerisce un impegno progressivo e sostenuto nella conservazione delle aree terrestri.

La **Germania** mostra un aumento significativo, passando dal 16% al 36%. Un andamento costantemente crescente riflette probabilmente una politica ambientale forte e coerente, volta a espandere e proteggere le aree naturali.

La **Spagna** registra un incremento da 14% a 27%, indicando un impegno crescente nel corso del tempo per la salvaguardia degli habitat naturali e la biodiversità.

I **Paesi Bassi** mostrano un incremento notevole, passando dal 4% a oltre il 20%. L'andamento è caratterizzato da periodi di stallo alternati a fasi di crescita, che potrebbero riflettere specifiche politiche o iniziative di conservazione.

Il **Belgio** ha un aumento moderato, da 10% a 15%, con periodi di stallo seguiti da crescita. Questo potrebbe indicare un approccio più graduale e misurato nella designazione di nuove aree protette.

La **Francia** mostra un incremento da 13% a 25%, con un andamento generalmente crescente. Questo suggerisce un impegno costante e in aumento nella conservazione delle aree terrestri.

In generale, i paesi dell'**Europa Occidentale** hanno mostrato una tendenza positiva nell'ampliamento delle aree terrestri protette, sebbene con ritmi e modalità diverse. Questo riflette una crescente consapevolezza dell'importanza della conservazione terrestre e della biodiversità in questa regione.

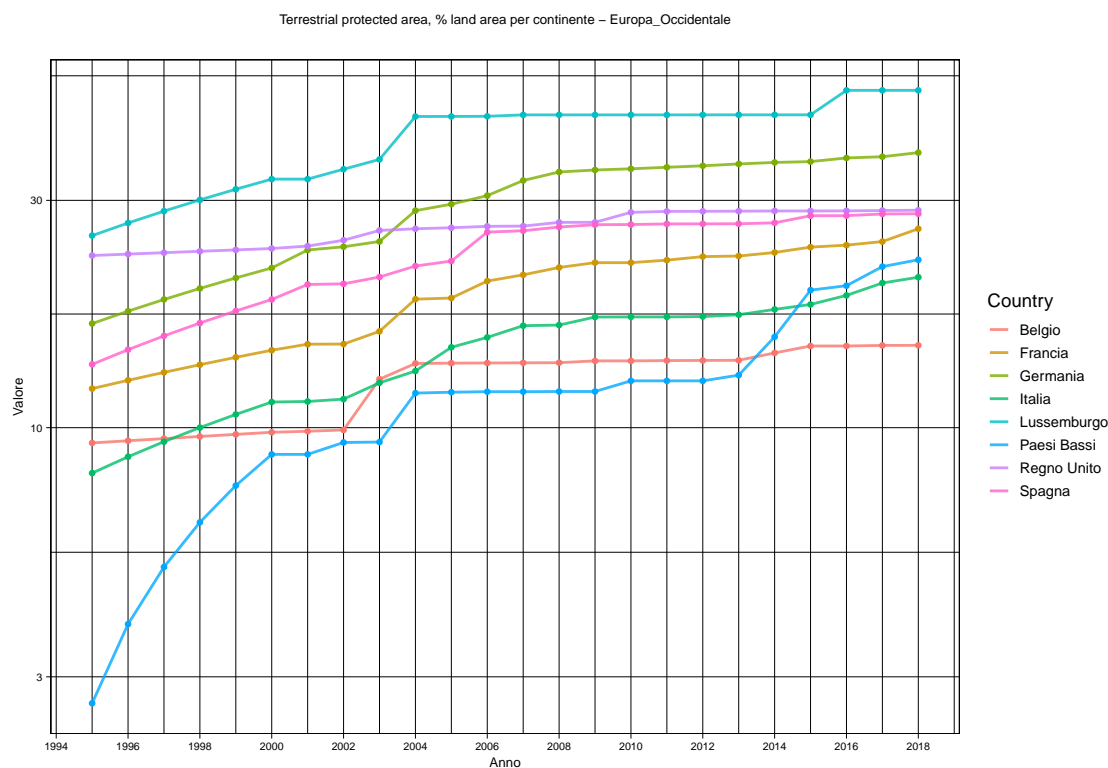


Figura 3.86: Serie storica per i paesi dell'Europa Occidentale

La figura 3.87 offre una panoramica dettagliata dell'evoluzione delle aree terrestri protette nei

paesi dell'Europa Orientale, mettendo in luce la diversità degli approcci e l'entità degli sforzi di conservazione nel corso degli anni.

La **Polonia** mostra un aumento significativo delle aree terrestri protette, passando dal 21% al 40%. Il raggiungimento del picco nel 2011, seguito da periodi di stallo, potrebbe indicare un impegno iniziale forte nella designazione di nuove aree protette, seguito da una fase di consolidamento.

In **Germania** si osserva un incremento da 10% a 35%, con un picco raggiunto nel 2009. Dopo questo periodo, sembra che ci sia stato un rallentamento o una stabilizzazione. Questo andamento potrebbe riflettere una politica ambientale inizialmente aggressiva, seguita da un periodo di mantenimento delle aree protette esistenti.

La **Danimarca** si mantiene relativamente costante attorno al 15%, suggerendo un approccio più uniforme e meno dinamico alla conservazione delle aree terrestri.

Il **Portogallo** mostra un incremento notevole, da 8% a 22%. Questa crescita potrebbe essere il risultato di politiche ambientali mirate e di un impegno crescente nel proteggere gli ecosistemi naturali. L'Ungheria, partendo da un livello leggermente più basso rispetto al Portogallo, segue un andamento simile, sovrapponendosi in alcuni punti. Questo suggerisce che entrambi i paesi potrebbero aver adottato approcci simili o aver risposto a iniziative ambientali comuni.

L'**Italia**, iniziando dallo stesso livello della Germania, mostra un andamento che si sovrappone alla Svezia, raggiungendo circa il 14%. Questo suggerisce un impegno moderato nella conservazione terrestre, potenzialmente influenzato da fattori come la disponibilità di aree naturali idonee e le priorità politiche.

Svezia e Finlandia mostrano aumenti più modesti, con la Svezia partendo dal 4% e la Finlandia da meno del 10%. Entrambi i paesi raggiungono livelli intorno al 13%, con vari stalli nel tempo. Questo andamento potrebbe riflettere una crescita graduale e mirata nella conservazione delle aree terrestri.

La figura 3.88 illustra l'evoluzione delle aree terrestri protette in **Australia, Nuova Zelanda e Sud Africa** hanno gestito la conservazione del loro territorio nel tempo.

In **Nuova Zelanda**, si assiste a un incremento costante e sostenuto della percentuale di aree terrestri protette, passando dal 25% a oltre il 35%. L'assenza di stalli in questo percorso suggerisce un impegno coerente e continuo nel proteggere e conservare gli ambienti naturali.

Il **Sud Africa** mostra una crescita più lenta e meno costante, con vari stalli, passando solo dal 5% a circa il 6%. Questo ritmo più graduale potrebbe indicare sfide nel bilanciare le esigenze di conservazione con altri obiettivi di sviluppo.

L'**Australia**, partendo da un livello simile al Sud Africa, mostra un incremento costante, raggiungendo circa il 20%. Questa crescita sostenuta potrebbe essere il risultato di politiche ambientali proattive e di un impegno nazionale nella protezione di ecosistemi unici come foreste, deserti e aree costiere.

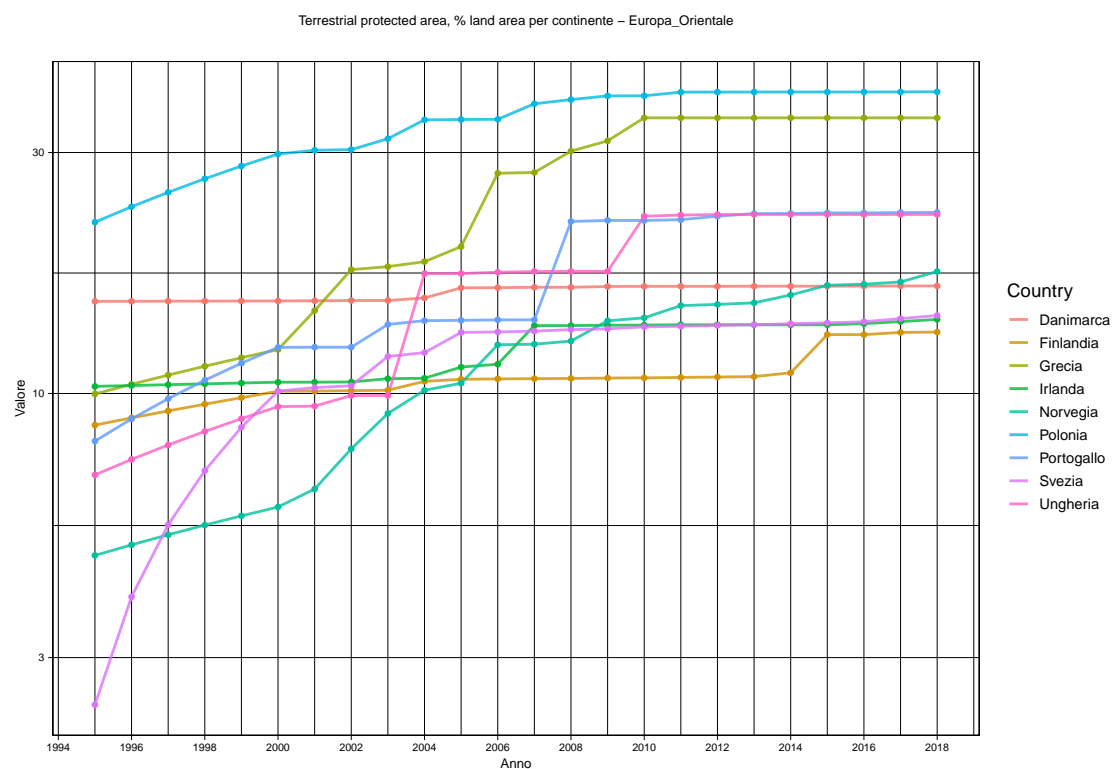


Figura 3.87: Serie storica per i paesi dell'Europa Orientale

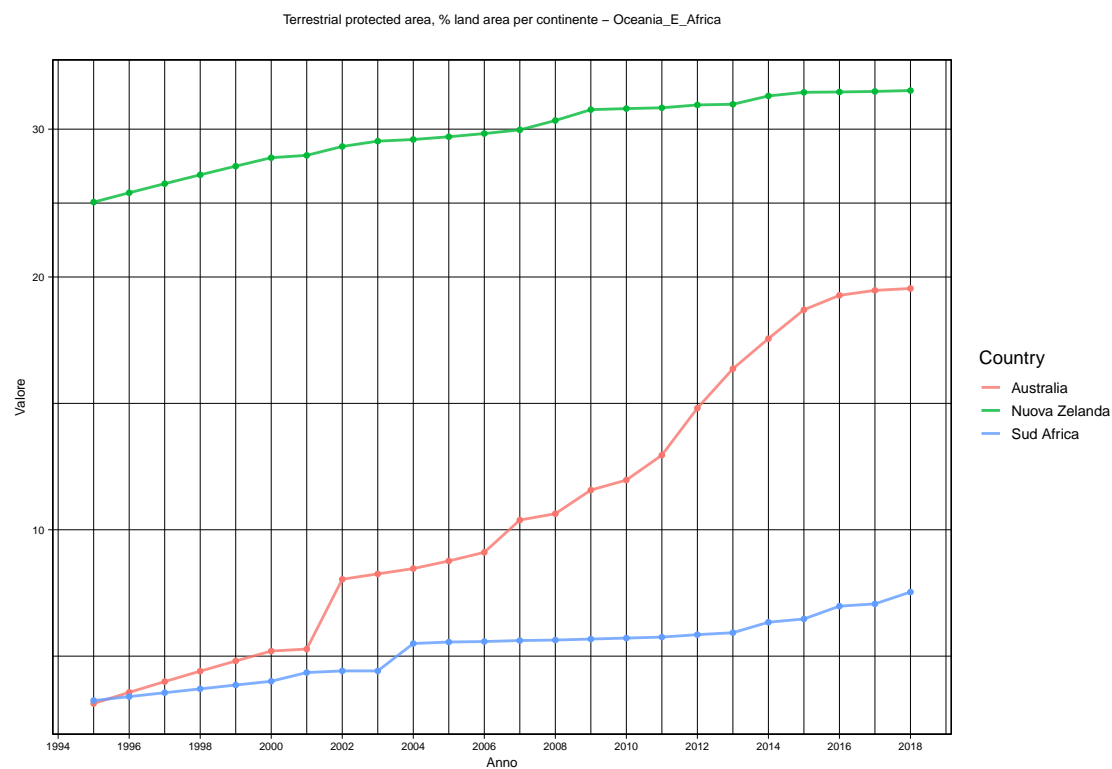


Figura 3.88: Serie storica per i paesi dell'Oceania e dell'Africa

3.2.12 Tasse Correlate all' Ambiente, % del Prodotto Interno Lordo (PIL)

Il gettito fiscale correlato all'ambiente è espresso come percentuale Prodotto Interno Lordo (PIL). Le tasse ambientali correlate includono

- prodotti energetici per scopi di trasporto (benzina e diesel) e per scopi stazionari (combustibili fossili ed elettricità);
- veicoli a motore e trasporti (imposte una tantum sull'importazione o vendita, imposte periodiche sulla registrazione o sull'uso stradale e altre imposte sui trasporti);
- gestione dei rifiuti (smaltimento finale, imballaggi e altre imposte sui prodotti legati ai rifiuti)
- sostanze che danneggiano lo strato di ozono e altre tasse correlate all'ambiente.

I dati fiscali sono ottenuti dal dataset degli strumenti utilizzati per la politica ambientale del Database sull'Ambiente dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE).

Tasse Correlate all'Ambiente, % del Prodotto Interno Lordo (PIL)	
Media	2.12
Mediana	2.24
Simmetria	0.1186233
Curtosi	0.6690066
Frequenza assoluta	7444
Massimo	5.36
Minimo	-1.53
Deviazione standard	0.9426049
Coefficiente di variazione	0.4445899
Varianza	0.888504
Q1	1.38
Q3	2.67

Tabella 3.12: Statistiche descrittive per la variabile: Tasse Correlate all'Ambiente, % del Prodotto Interno Lordo (PIL)

La **Media** (2.12%) indica che, in media, le tasse correlate all'ambiente rappresentano poco più del 2% del Prodotto Interno Lordo (PIL) nei paesi inclusi nello studio. Questo riflette un impegno moderato nella tassazione ambientale a livello globale. La **Mediana (2,24%)** leggermente superiore alla media, indicando che più della metà dei paesi hanno una quota di tasse ambientali superiore alla media, mostrando una generale inclinazione verso la tassazione ambientale.

Una **simmetria quasi neutra** suggerisce una distribuzione quasi simmetrica dei dati intorno alla media. La **curtosi** indica una distribuzione leggermente più appuntita rispetto a una distribuzione normale, ma non in modo significativo. Con un **massimo** del 5.36% e un **minimo** sorprendentemente negativo di -1.53%, l'intervallo suggerisce una varietà di approcci alla tassazione ambientale tra i paesi. La **Deviazione Standard** (0.9426049) indica una variazione moderata nei livelli di tassazione ambientale tra i paesi. La **varianza** (0.888504) conferma la variabilità moderata evidenziata dalla deviazione standard. I quantili Q1 a 1.38% e Q3 a 2.67% mostrano che la metà dei paesi ha valori di tassazione ambientale compresi in questo intervallo.

Il **boxplot** 3.89) mostra una mediana leggermente superiore alla media, con una distribuzione relativamente equilibrata dei dati. La presenza di valori estremi (outliers), come indicato dal valore minimo negativo, potrebbe suggerire alcune anomalie o particolari condizioni politiche ed economiche in alcuni paesi.

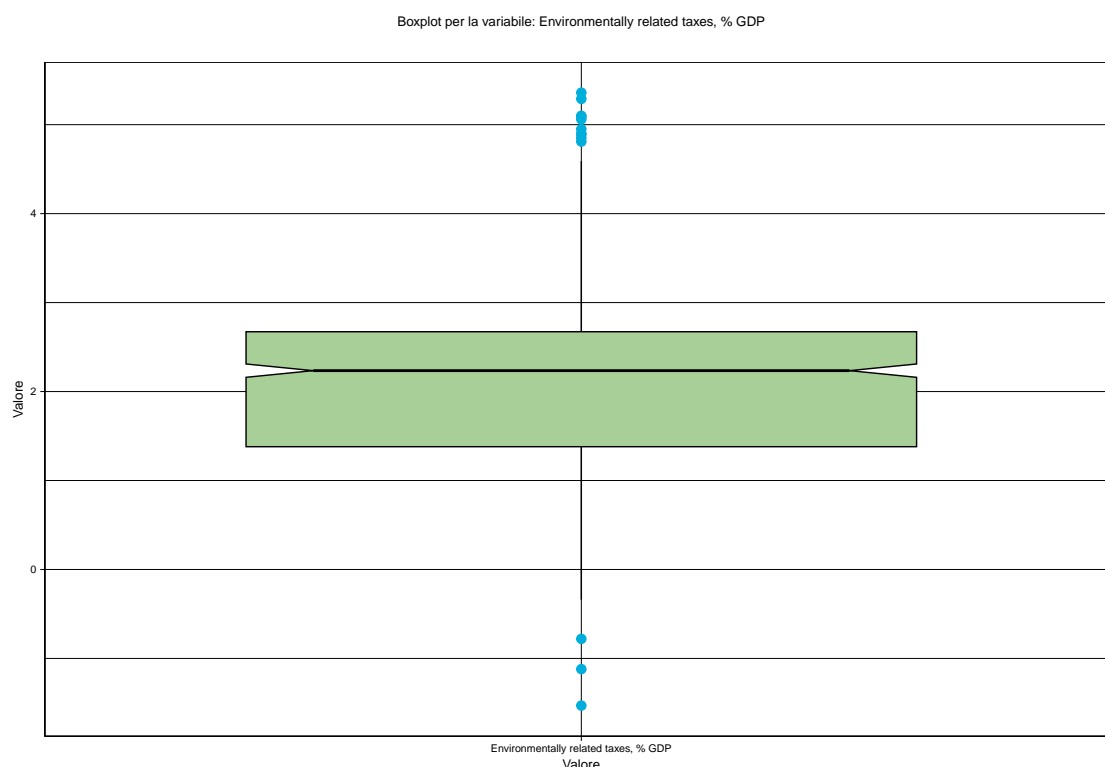


Figura 3.89: Boxplot per la variabile: Tasse Correlate all'Ambiente, % del Prodotto Interno Lordo (PIL)

L'**istogramma** 3.90 mostra una concentrazione di paesi con valori di tassazione ambientale inferiori alla media, con una distribuzione che si estende verso valori più alti. La presenza di una coda verso i valori più alti suggerisce che alcuni paesi hanno adottato un approccio più aggressivo alla tassazione ambientale.

La **funzione di densità** 3.91 conferma la distribuzione osservata nell'istogramma, con un picco nei valori più bassi e una coda che si estende verso i valori più alti.

La figura 3.92 illustra la variazione delle tasse correlate all'ambiente come percentuale del Prodotto Interno Lordo (PIL) nei paesi del **continente americano** nel periodo analizzato.

Il **Messico** mostra una variazione notevole nel tempo. Dopo aver raggiunto il picco di oltre 1.5% nel 1988, segue un periodo di fluttuazione con alti e bassi, culminando in un valore negativo di -1.53%, un indicatore insolito che potrebbe riflettere specifiche circostanze politiche o economiche durante quel periodo. Dopo il 2012, osserviamo un recupero fino a 1.5% nel 2016, seguito da una nuova discesa a 1% nel 2018. Questo andamento potrebbe indicare cambiamenti nelle politiche

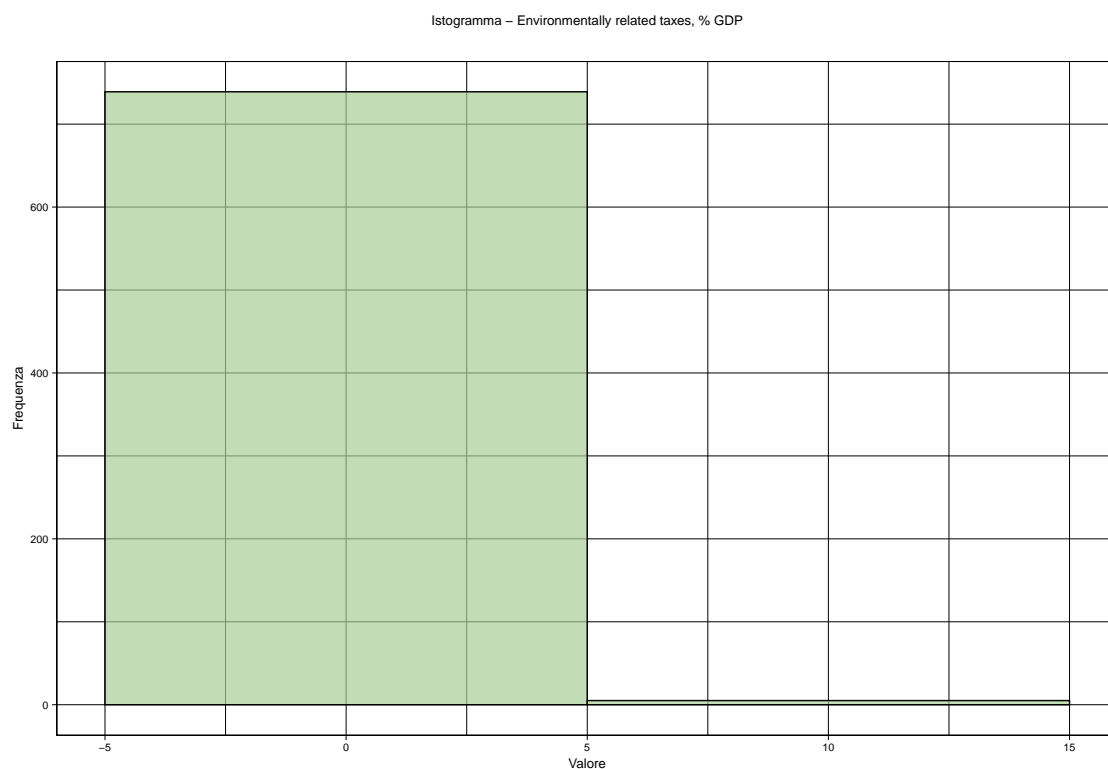


Figura 3.90: Istogramma

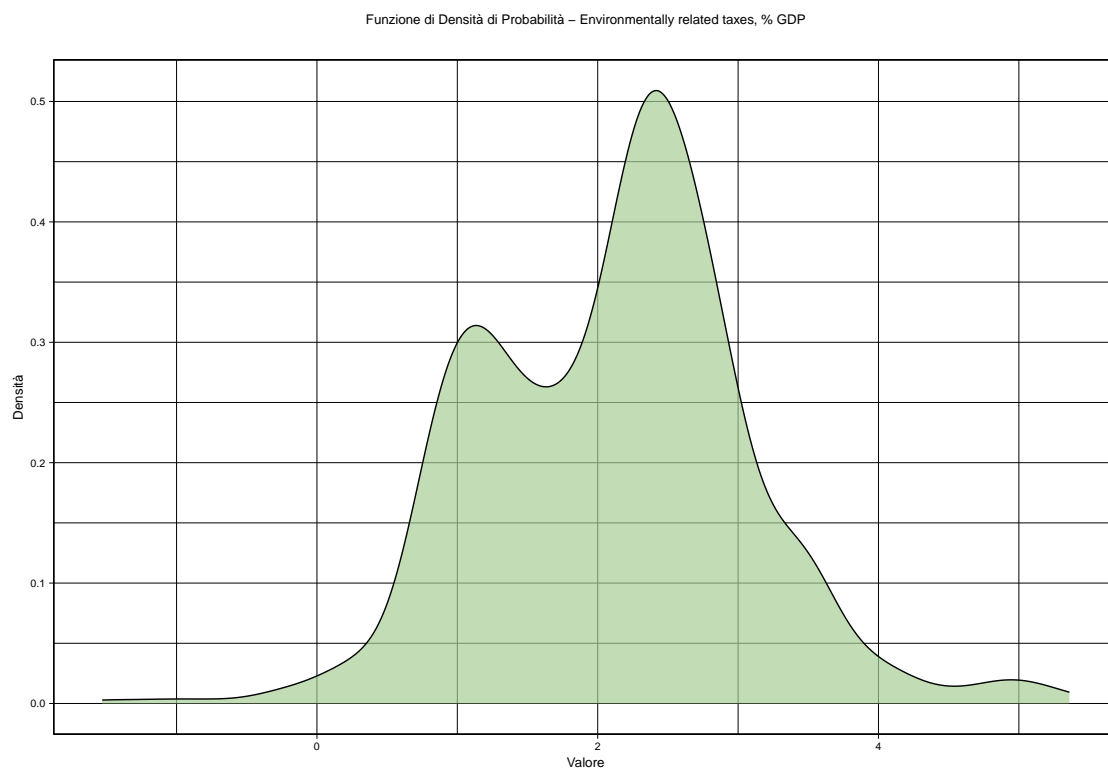


Figura 3.91: Funzione di densità

ambientali o economiche nel corso degli anni.

Il **Brasile** inizia con valori vicino allo zero e aumenta gradualmente, sebbene in modo limitato, per arrivare a meno di 1%. Questo suggerisce un aumento contenuto nella tassazione ambientale, forse riflettendo un approccio graduale o limitato all'imposizione di tasse ambientali.

Il **Canada** mantiene una consistenza relativa, oscillando tra 1.5% e 1%. Questo suggerisce una certa stabilità nelle politiche ambientali, con una tassazione ambientale che rimane all'interno di un intervallo ristretto.

L'**Argentina** mostra un'oscillazione significativa. Dopo un aumento fino a quasi 3% nel 1998, seguono periodi di incremento e diminuzione, concludendo a meno di 2% nel 2018. Questo potrebbe riflettere cambiamenti nella politica ambientale o variazioni nell'economia.

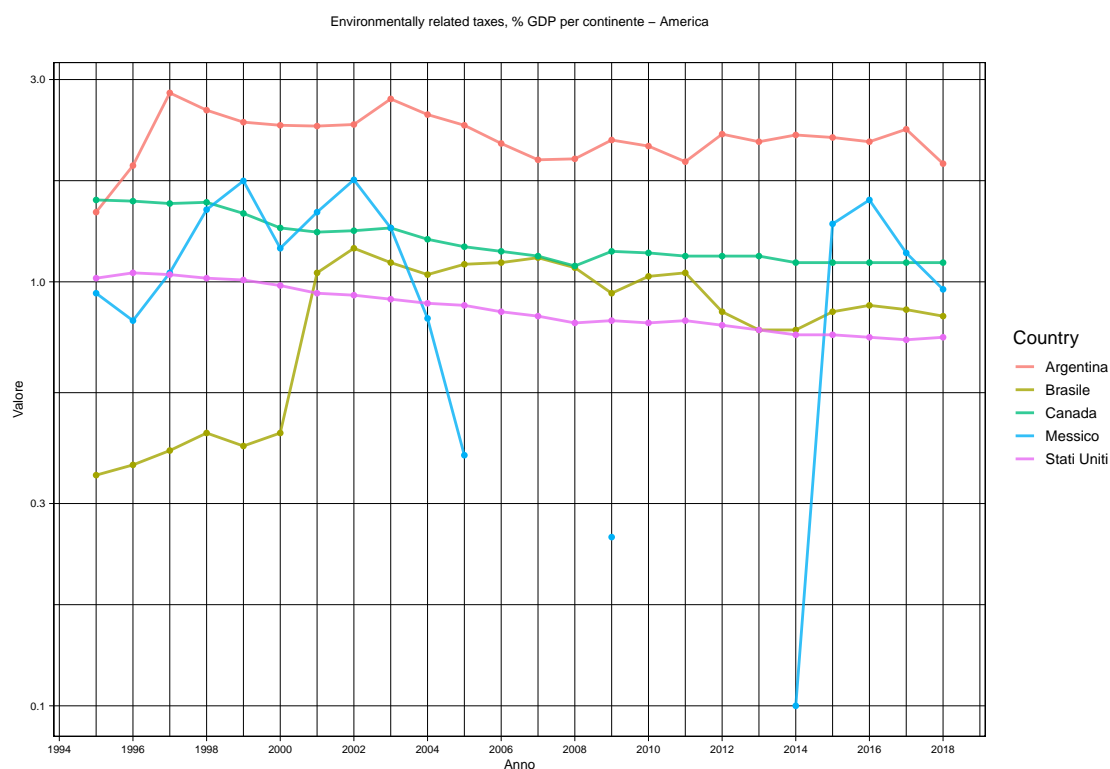


Figura 3.92: Serie storica dei paesi del continente Americano

Nella figura 3.93, osserviamo le variazioni delle tasse correlate all'ambiente come percentuale del Prodotto Interno Lordo (PIL) nei **paesi asiatici** selezionati.

Russia e Indonesia mostrano una relativa stabilità nel tempo con valori oscillanti intorno a 1%, suggerendo una tassazione ambientale consistente e una politica fiscale ambientale relativamente stabile. La mancanza di variazioni significative potrebbe indicare un approccio uniforme alla tassazione ambientale nel corso degli anni.

La **Corea del Sud** parte da un valore inferiore a 2%, poi mostra una tendenza crescente, sebbene con fluttuazioni. Questi picchi di crescita e diminuzione potrebbero riflettere cambiamenti nelle politiche ambientali o economiche, con un incremento generale della tassazione ambientale nel tempo. Questo potrebbe essere interpretato come un impegno crescente nella gestione delle questioni ambientali.

In **Giappone** si osserva una lieve diminuzione da poco sopra 1.5% a sotto 1.5%. Questo calo graduale potrebbe indicare un leggero allentamento della pressione fiscale ambientale o cambiamenti nell'economia che influenzano la percentuale del Prodotto Interno Lordo (PIL) attribuita a tali tasse.

La **Cina** parte da un valore basso, vicino allo zero, e raggiunge il picco di oltre 1.5% nel 2012. Successivamente, si nota una diminuzione fino a poco sotto 1%. Questa variazione può riflettere cambiamenti sostanziali nelle politiche ambientali e economiche del paese, con un incremento iniziale seguito da una riduzione, possibilmente legata a trasformazioni nel contesto economico o nelle priorità politiche.

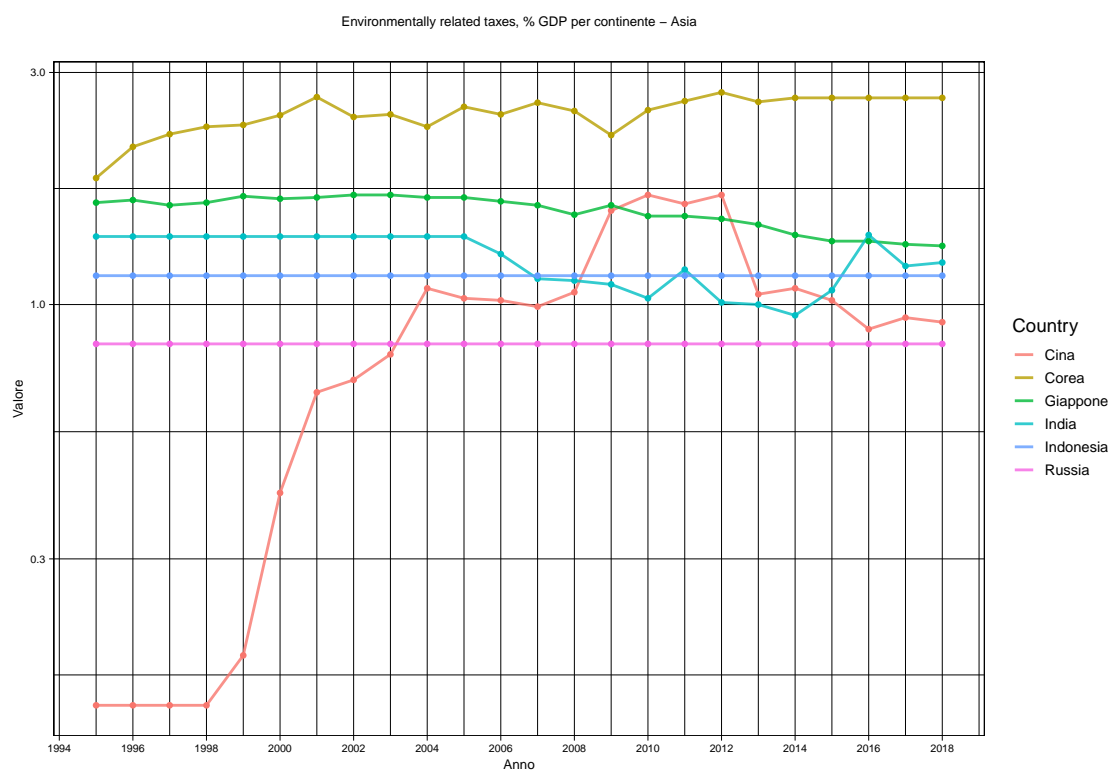


Figura 3.93: Serie storica dei paesi del continente Asiatico

La **Danimarca** evidenzia un andamento altalenante nel corso degli anni. All'inizio del periodo osservato, il valore è inferiore a 4. Nel 1999, il valore raggiunge un picco di oltre 5, per poi seguire una traiettoria di diminuzione e successiva risalita negli anni successivi. Questa variabilità porta il valore a stabilizzarsi attorno a 3.5 entro il 2018.

Per quanto riguarda gli altri paesi analizzati, le tendenze sono generalmente simili, con valori che oscillano tra 2.5 e 3. Questi valori si mantengono relativamente costanti nel tempo, senza mostrare

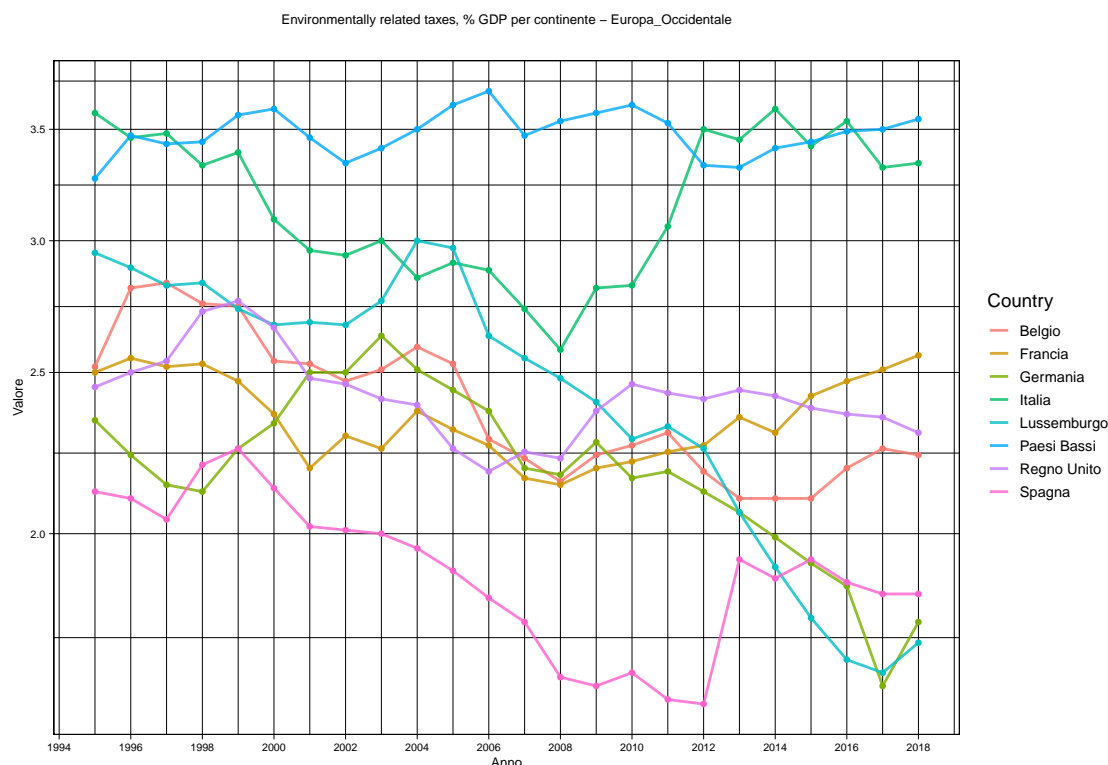


Figura 3.94: Serie storica dei paesi dell'Europa Occidentale

significative variazioni. Tuttavia, è possibile notare una certa variabilità nel caso della **Grecia**. In particolare, il valore iniziale è pari a 3, per poi scendere fino a 2. Successivamente, si osserva una ripresa che porta il valore a risalire fino a 4.

Nella figura 3.96, l'andamento delle tasse correlate all'ambiente come percentuale del Prodotto Interno Lordo (PIL) in **Australia**, **Nuova Zelanda** e **Sud Africa** offre una panoramica interessante.

L'**Australia** mostra un andamento complessivamente decrescente, partendo da 2.5% e arrivando a 0.4%. Tuttavia, questo percorso è caratterizzato da notevoli fluttuazioni, con picchi e cali che potrebbero indicare cambiamenti nelle politiche fiscali o ambientali o variazioni nella composizione economica del paese.

La **Nuova Zelanda** parte da 1.5% e scende leggermente a poco sopra 1.25%, anch'essa mostrando una serie di alti e bassi. Questa fluttuazione potrebbe riflettere adattamenti o ricalibramenti periodiche delle politiche fiscali ambientali. La relativa stabilità del valore finale suggerisce che, nonostante le variazioni, la tassazione ambientale rimane una componente consistente del sistema fiscale del paese.

Il **Sud Africa** parte da 1.9% e aumenta fino a 3%, sebbene con un andamento non lineare. Questo incremento, intervallato da periodi di variazione, indica un possibile rafforzamento delle politiche ambientali attraverso misure fiscali.

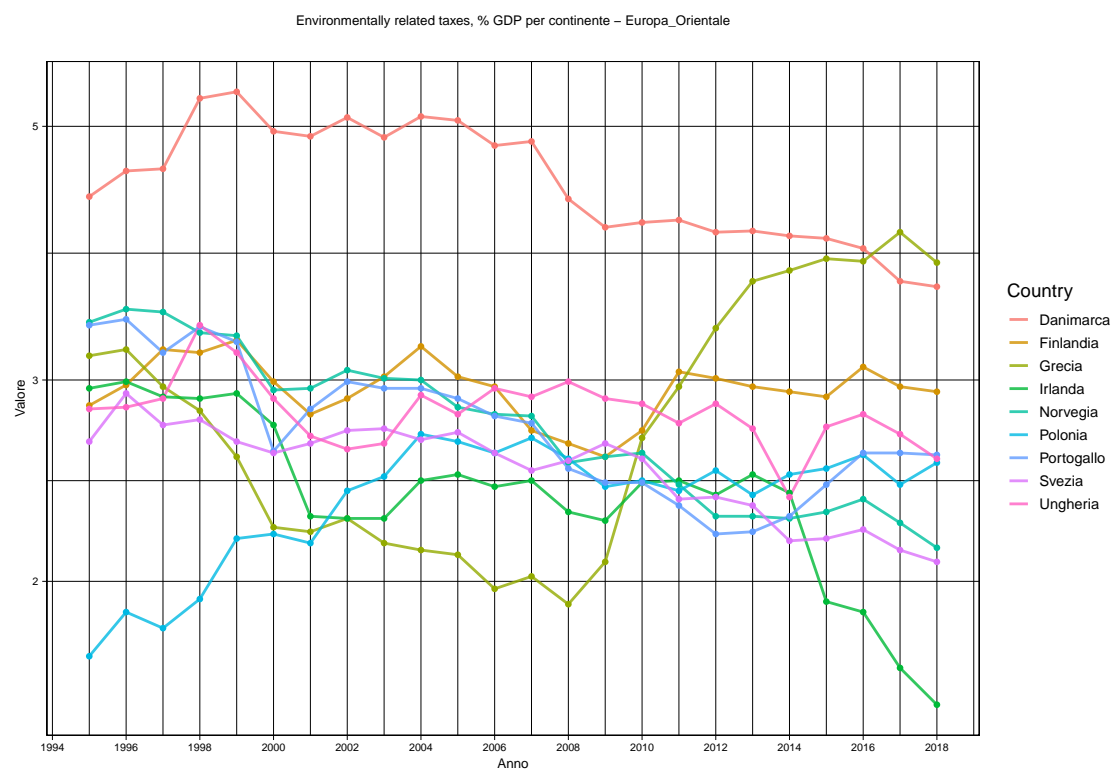


Figura 3.95: Serie storica dei paesi dell'Europa Orientale

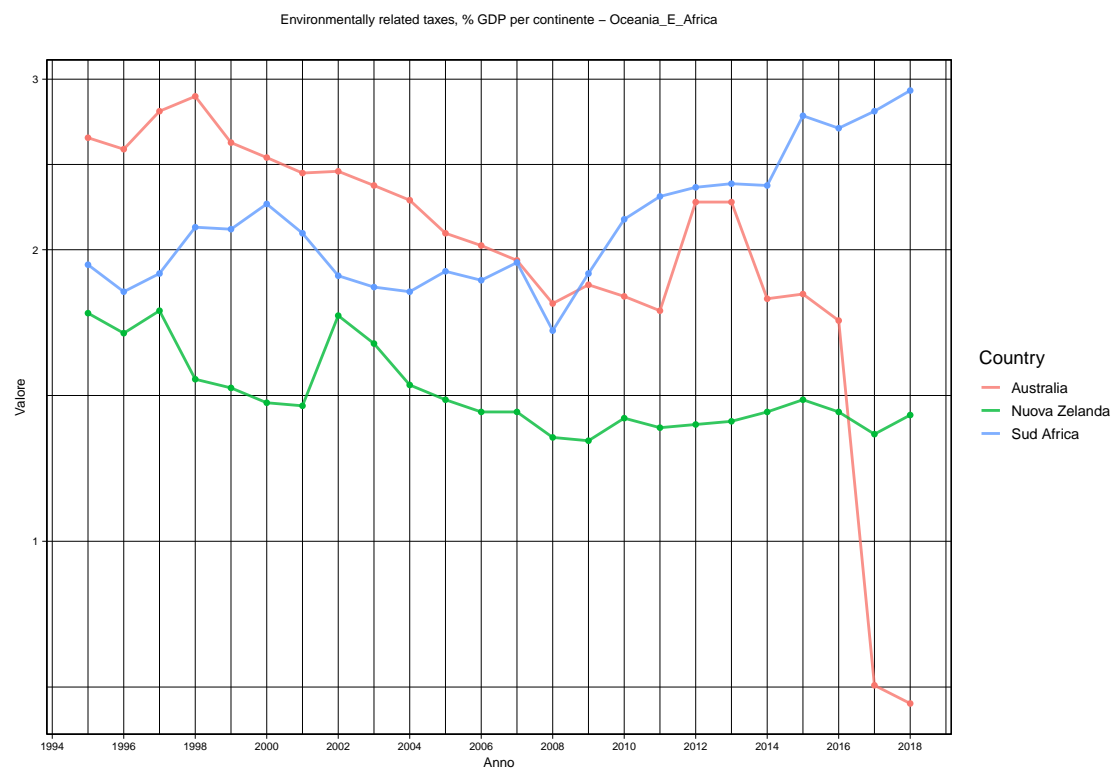


Figura 3.96: Serie storica per i paesi dell'Oceania e dell'Africa

3.2.13 Prodotto Interno Lordo (PIL)

Il Prodotto Interno Lordo (PIL) pro capite in dollari statunitensi costanti del 2015, espressi a parità di potere d'acquisto. Questo indicatore misura la ricchezza economica di una nazione divisa per la sua popolazione. Tuttavia, poiché si tratta di un valore medio, non riflette la distribuzione del reddito. Inoltre, è una misura "lorda" del reddito e non tiene conto né dell'ammortamento dei beni prodotti né dell'esaurimento delle risorse naturali.

Prodotto Interno Lordo (PIL)	
Media	1 34175.
Mediana	35787.
Simmetria	1.141464
Curtosi	3.167752
Frequenza assoluta	7444
Massimo	114974.
Minimo	1931
Deviazione standard	19347.
Coefficiente di variazione	0.5661222
Varianza	374318480
Q1	18909.
Q3	43886.

Tabella 3.13: Statistiche descrittive per la variabile:Prodotto Interno Lordo (PIL)

La **media** del Prodotto Interno Lordo (PIL) pro capite è di 34.175 USD, mentre la **mediana** è di 35.787 USD. La mediana leggermente superiore alla media suggerisce che più della metà dei paesi ha un Prodotto Interno Lordo (PIL) pro capite superiore alla media, indicando una distribuzione asimmetrica con una coda verso i valori più bassi.

Un **coefficiente di simmetria** di 1.141464 indica una distribuzione asimmetrica positiva, con una coda che si estende verso i valori più alti. La **curtosi** di 3.167752 suggerisce una distribuzione più appuntita rispetto a una normale, indicando una maggiore frequenza di valori estremi. La **deviazione standard** di 19.347 USD e una varianza di circa 374 milioni riflettono una variabilità significativa nel Prodotto Interno Lordo (PIL) pro capite tra i paesi. Questo suggerisce notevoli differenze nella ricchezza economica per persona a livello globale.

Il Q1 a 18.909 USD e il Q3 a 43.886 USD mostrano che la metà dei paesi ha un Prodotto Interno Lordo (PIL) pro capite compreso in questo intervallo. L'ampio range tra il **minimo** di 1.931 United States Dollar (USD) e il **massimo** di 114.974 USD sottolinea l'estrema disparità nella ricchezza economica per persona tra i paesi più poveri e quelli più ricchi.

Il **boxplot** mostra una mediana vicino al limite superiore della scatola, indicando una concentrazione di paesi con PIL pro capite più alto. La presenza di valori anomali suggerisce che alcuni paesi hanno un PIL pro capite eccezionalmente alto.

L'**istogramma** mostra una distribuzione inclinata verso destra, con una maggioranza di paesi che hanno un Prodotto Interno Lordo (PIL) pro capite inferiore alla media.

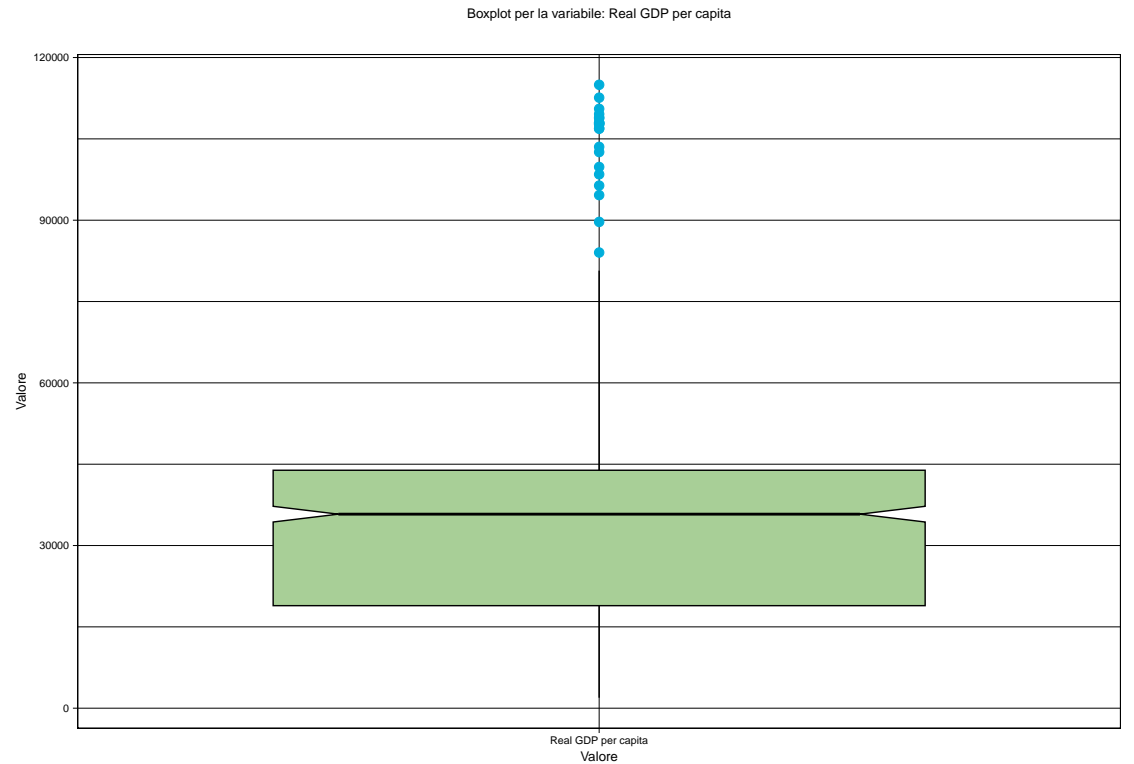


Figura 3.97: Boxplot per la variabile: Prodotto Interno Lordo (PIL)

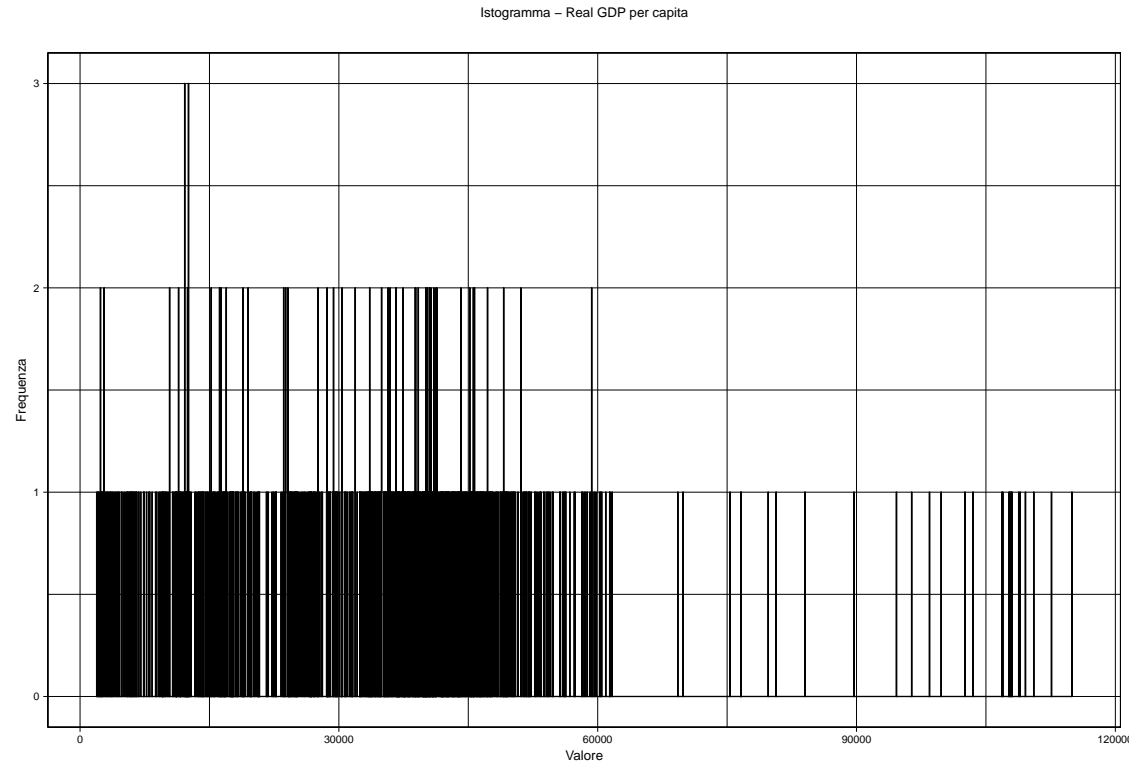


Figura 3.98: Istogramma

La **funzione di densità** conferma l'asimmetria positiva e la curtosi osservata, con un picco nei valori più bassi e una coda lunga verso i valori più alti.

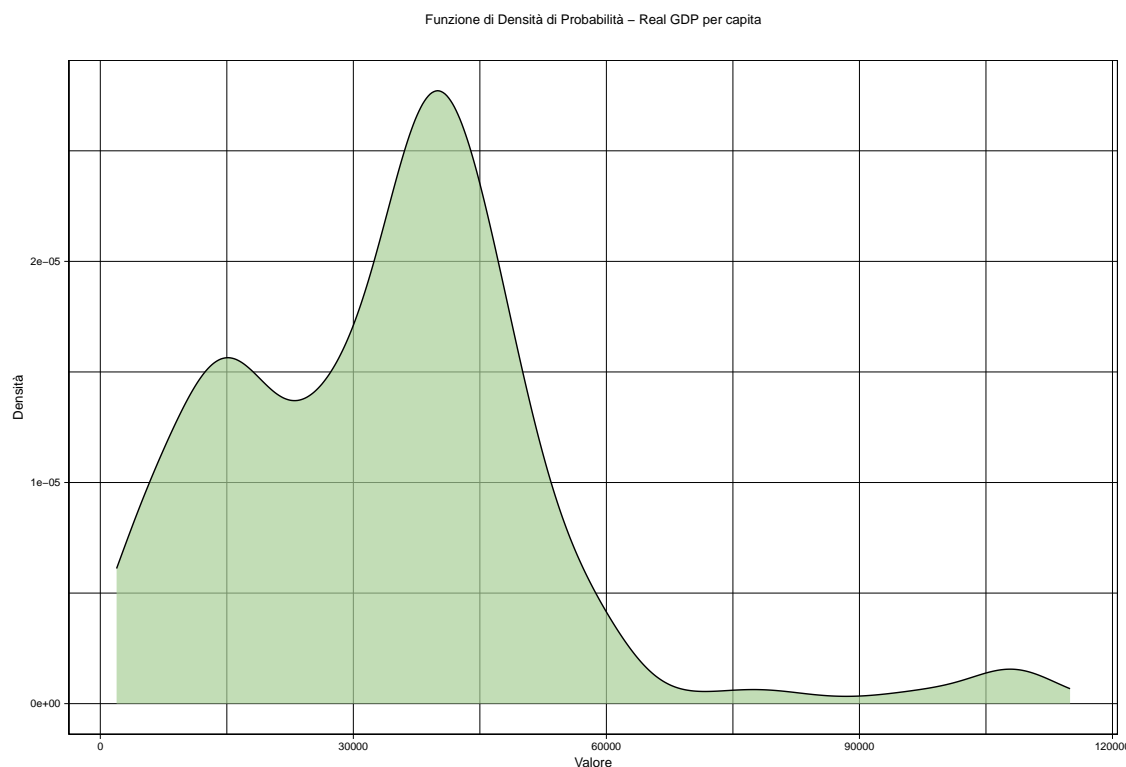


Figura 3.99: Funzione di densità

In sintesi, i dati riflettono una significativa disparità economica globale, con alcuni paesi che godono di un elevato Prodotto Interno Lordo (PIL) pro capite mentre molti altri rimangono significativamente più indietro. Questa distribuzione può avere implicazioni per le politiche di sviluppo, l'assistenza internazionale e le strategie di riduzione della povertà.

Nella figura 3.100, viene illustrato l'andamento del PIL pro capite per diversi paesi del **continente americano**. Il grafico mostra un aumento generale del PIL pro capite nel tempo per la maggior parte dei paesi analizzati. Gli **Stati Uniti** si distinguono con valori significativamente più alti rispetto agli altri paesi, mantenendo una traiettoria di crescita relativamente stabile. L'**Argentina** evidenzia un andamento più altalenante, con un picco attorno al 2008, seguito da un declino e una successiva ripresa. Il **Messico** mostra un andamento variabile, con forti fluttuazioni che potrebbero riflettere le crisi economiche che hanno colpito il paese. Il **Canada** e il **Brasile** mostrano tendenze di crescita più moderate e stabili nel tempo.

La figura 3.101 mostra la serie storica del PIL pro capite per i principali **paesi asiatici**. La **Cina** mostra una crescita impressionante, partendo da valori molto bassi per raggiungere un notevole incremento nel periodo analizzato, riflettendo la rapida industrializzazione ed espansione economica del paese. La **Corea del Sud** e il **Giappone** mostrano valori iniziali molto più alti, con la Corea che continua a crescere, mentre il Giappone mostra una crescita più contenuta e stabile. L'**India**, pur mostrando una crescita, rimane su valori inferiori rispetto agli altri paesi. La **Russia**

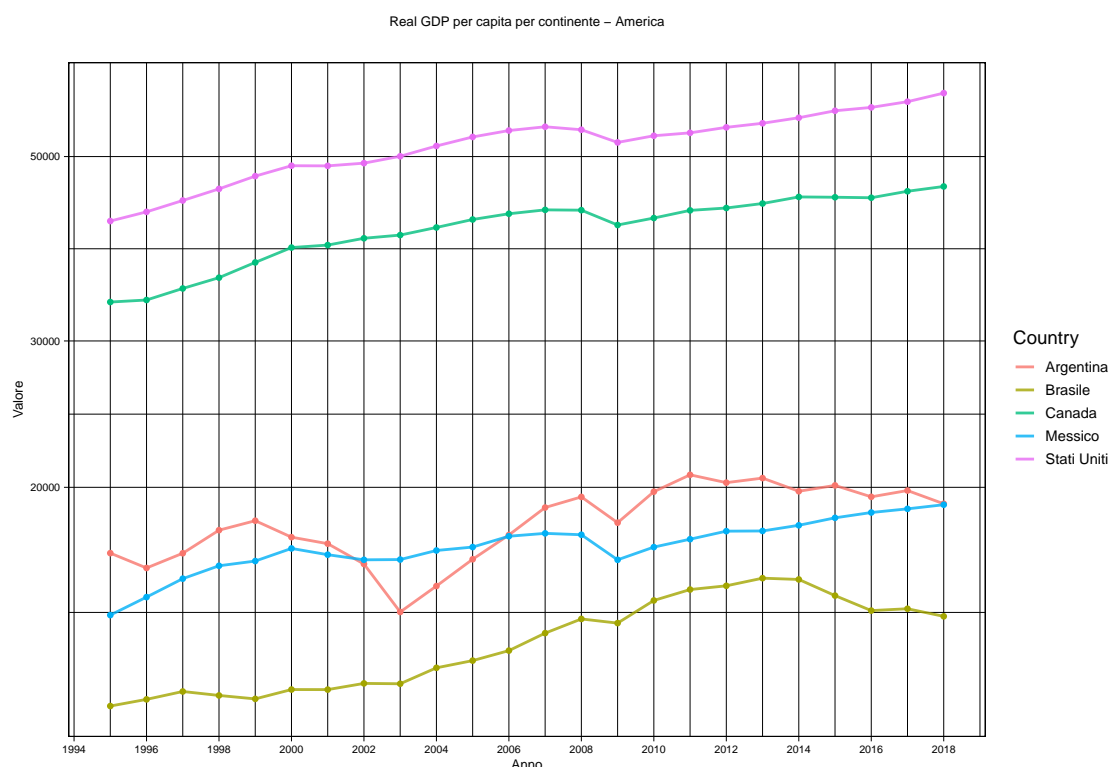


Figura 3.100: Serie storica per i paesi del continente americano, Prodotto Interno Lordo (PIL)

e l'**Indonesia** mostrano una crescita moderata, con la Russia che parte da valori relativamente alti e l'Indonesia da valori più bassi.

Nella figura 3.102, viene illustrata la variazione del PIL pro capite per i paesi dell'**Oceania** e del **Sud Africa**. L'**Australia** e la **Nuova Zelanda** mostrano un andamento del PIL pro capite simile, con una crescita costante e valori relativamente alti. In contrasto, il **Sud Africa** parte da valori più bassi e mostra una crescita più irregolare, con un certo miglioramento verso la fine del periodo analizzato .

La figura 3.103 presenta l'andamento del PIL pro capite nei paesi dell'Europa orientale. Il grafico mostra la **Danimarca**, la **Norvegia**, e la **Svezia** con valori elevati e una crescita stabile nel tempo. La **Polonia** e l'**Ungheria** partono da valori inferiori, ma mostrano una crescita costante, riducendo gradualmente il divario con i paesi dell'Europa settentrionale. La **Grecia** mostra un andamento altalenante, con una significativa riduzione del PIL pro capite durante la crisi economica, seguita da una parziale ripresa.

Nella figura 3.104, viene mostrata la serie storica del PIL pro capite per i paesi dell'Europa occidentale. Il **Lussemburgo** si distingue con valori estremamente elevati rispetto agli altri paesi, mantenendo una crescita sostenuta. La **Germania** e i **Paesi Bassi** mostrano un andamento simile, con valori elevati e una crescita stabile. La **Francia**, il **Regno Unito**, e l'**Italia** mostrano una crescita più moderata, con l'Italia che evidenzia una leggera flessione negli anni più recenti. La **Spagna** mostra una crescita costante, ma parte da valori inferiori rispetto agli altri paesi dell'Europa occidentale.

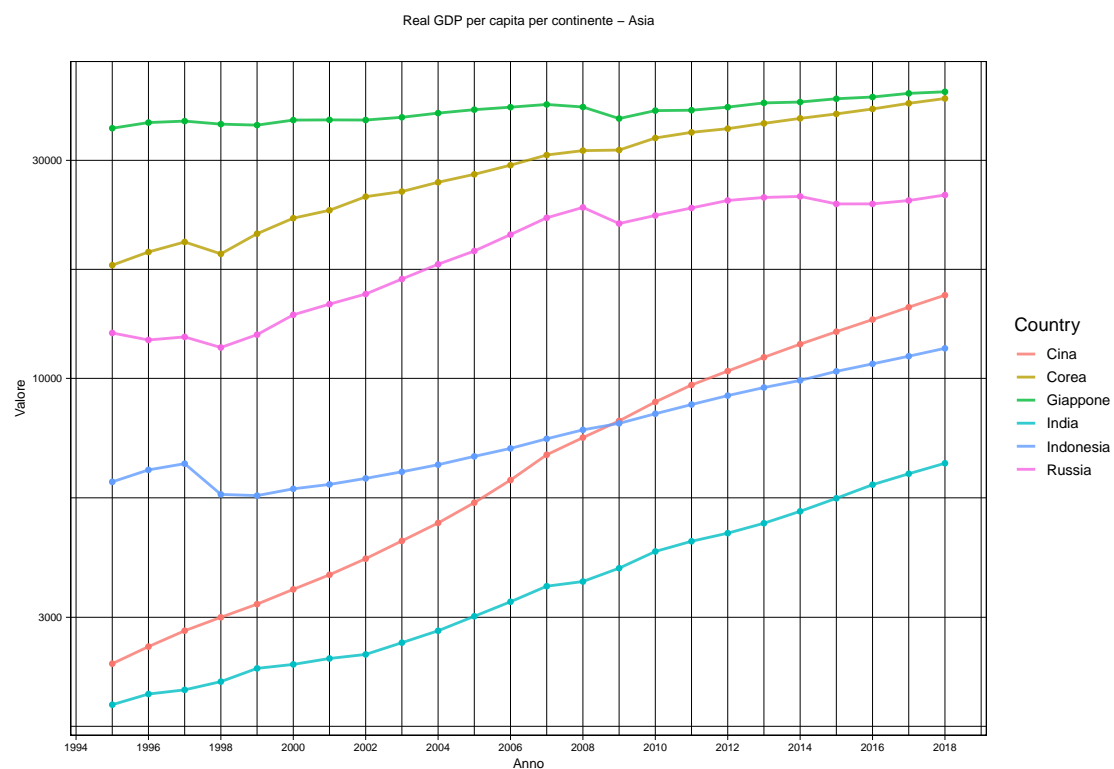


Figura 3.101: Serie storica per i paesi del continente Asiatico

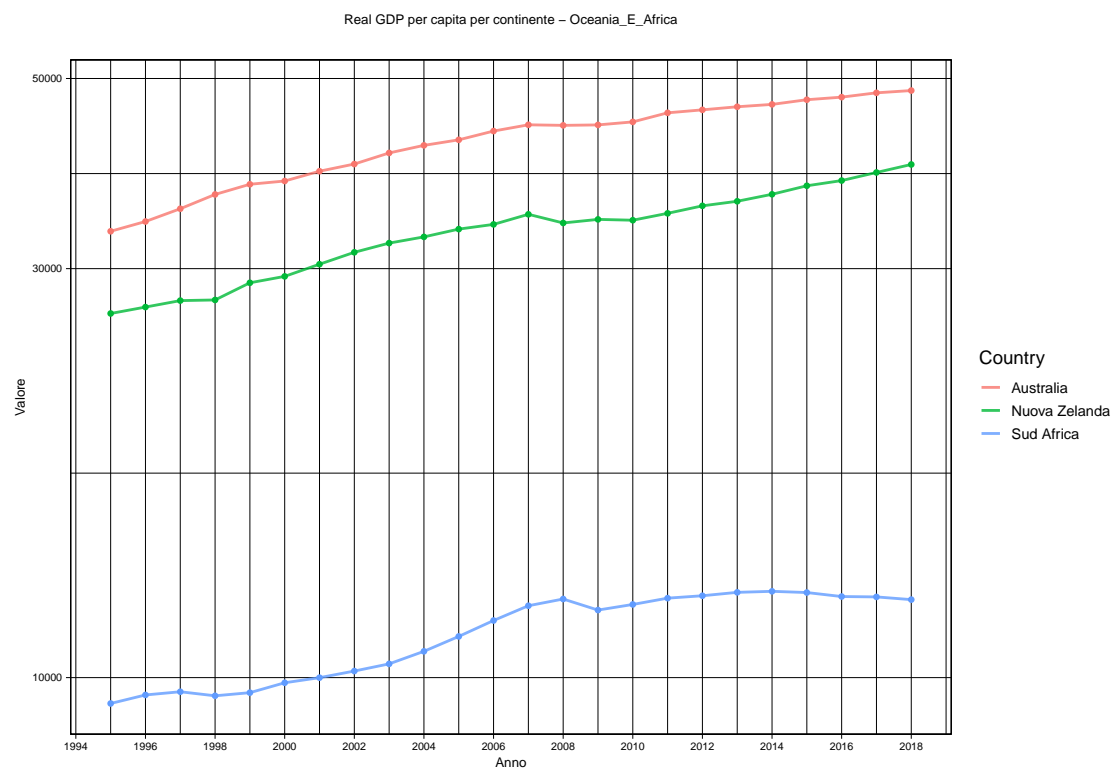


Figura 3.102: Serie storica per i paesi dell'Oceania e dell'Africa

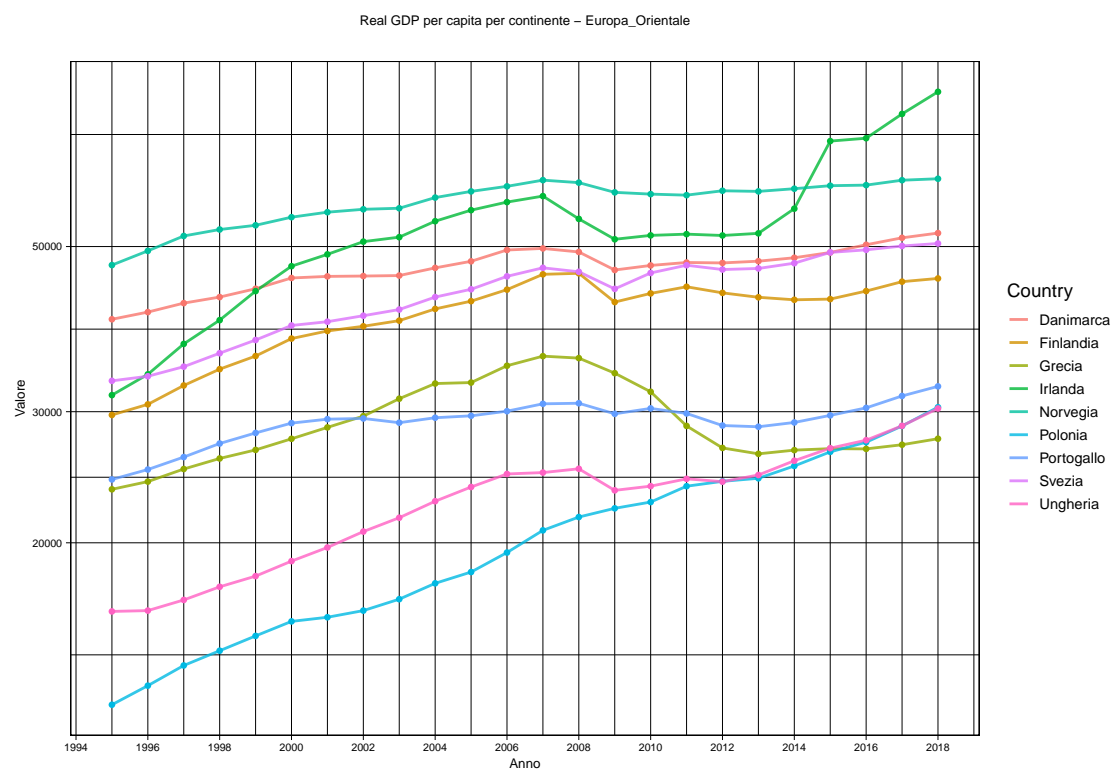


Figura 3.103: Serie storica per i paesi dell'Europa Orientale

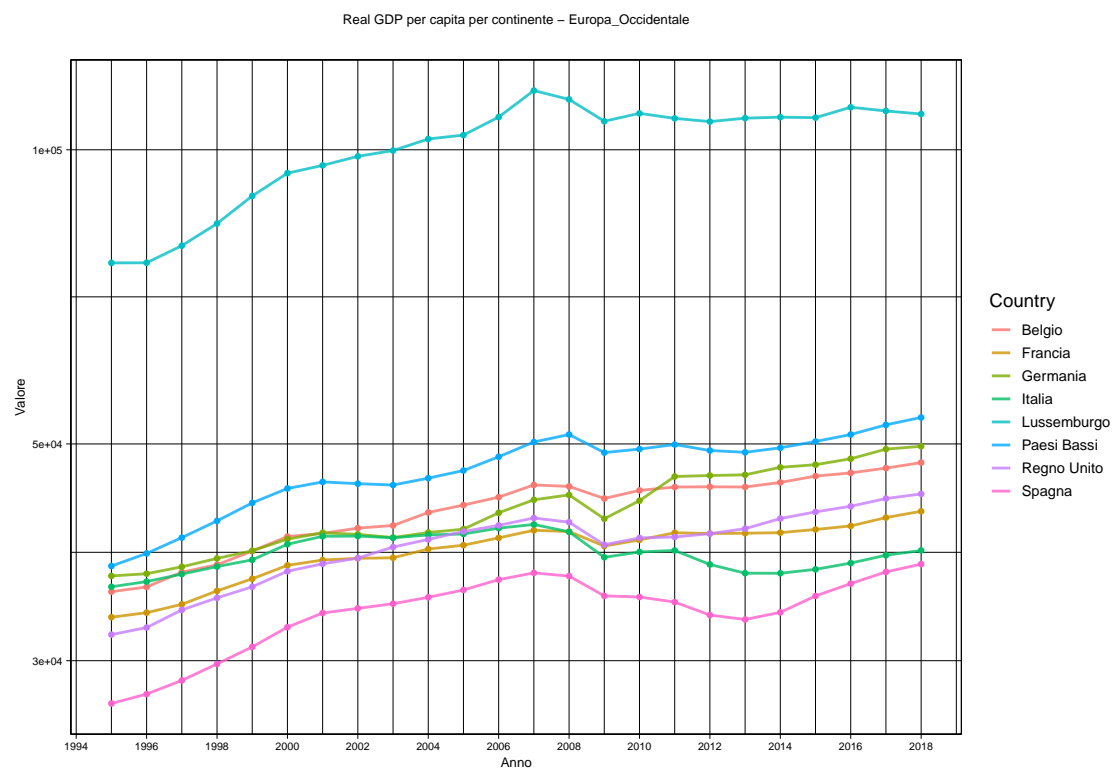


Figura 3.104: Serie storica per i paesi dell'Europa Occidentale

3.3 Diagramma di Pareto

Il diagramma di Pareto è uno strumento analitico fondamentale nel campo dell'analisi dei dati. Questo tipo di grafico è costituito da barre verticali ordinate in maniera decrescente per frequenza. Nel contesto di quest'analisi, i paesi sono rappresentati sull'asse delle x, mentre l'asse delle y quantifica variabili specifiche, come la frequenza di una determinata misura.

La **Legge di Pareto**, conosciuta anche come principio del 80/20, è una regola empirica che afferma che, per molti eventi, circa l'80% degli effetti proviene dal 20% delle cause. Questo principio, scoperto dall'economista italiano Vilfredo Pareto nel contesto della distribuzione della ricchezza, trova applicazione in numerosi ambiti, inclusi gli studi ambientali.

Nel contesto dell'analisi delle politiche ambientali, la **Legge di Pareto** può essere utilizzata per identificare un numero limitato di fattori chiave (il 20%) che sono responsabili della maggior parte (l'80%) degli impatti o dei risultati. Questo può rivelarsi particolarmente utile nelle seguenti aree:

1. **Investimenti in Energie Rinnovabili:** Pochi paesi potrebbero guidare la transizione globale verso le energie rinnovabili. Comprendere e supportare le politiche di questi paesi può avere un impatto ampio sulla sostenibilità energetica a livello mondiale.
2. **Gestione delle Risorse Idriche:** Un piccolo numero di nazioni potrebbe essere responsabile di una grande parte del prelievo idrico a livello mondiale. Identificare e intervenire su queste "cause principali" può portare a miglioramenti significativi nella gestione delle risorse idriche.
3. **Emissioni di CO₂:** Potrebbe emergere che una quota relativamente piccola di paesi sia responsabile della maggior parte delle emissioni globali di CO₂. Concentrando gli sforzi su questi pochi paesi, potrebbe essere possibile ottenere una riduzione significativa delle emissioni globali.

3.3.1 Investimenti in Energie Rinnovabili

L'analisi condotta attraverso il Diagramma di Pareto fornisce un'illuminante panoramica sull'utilizzo dell'energia rinnovabile a livello mondiale. Come evidenziato nel grafico (Figura 3.105), emerge un gruppo ristretto di paesi che si distinguono come principali produttori di energia rinnovabile. Tra questi, **Norvegia, Brasile, Indonesia, Nuova Zelanda, Svezia e India** si affermano per il loro significativo contributo, rappresentando insieme il **56%** dell'intera produzione di energia rinnovabile tra i 33 paesi analizzati. Questa osservazione è di particolare rilievo poiché sottolinea l'impegno di un numero limitato di nazioni nella guida della transizione globale verso fonti energetiche più sostenibili.

L'analisi complessiva indica che questi sei paesi, unitamente a **Finlandia, Portogallo, Danimarca, Canada, Cina, Italia e Messico**, contribuiscono per l'**80%** all'approvvigionamento totale di energia rinnovabile. Tale disuguaglianza evidenzia la disparità esistente tra i paesi pionieri in questo settore e quelli ancora in fase di sviluppo della propria infrastruttura di energie rinnovabili.

Come si può osservare nella figura 3.105, è evidente un marcato divario tra i paesi in testa alla classifica e gli altri. Questo aspetto riflette non solo le differenze nelle capacità e nelle politiche nazionali riguardanti le energie rinnovabili, ma anche l'esigenza di interventi mirati per promuovere

una distribuzione più equa delle risorse energetiche rinnovabili.

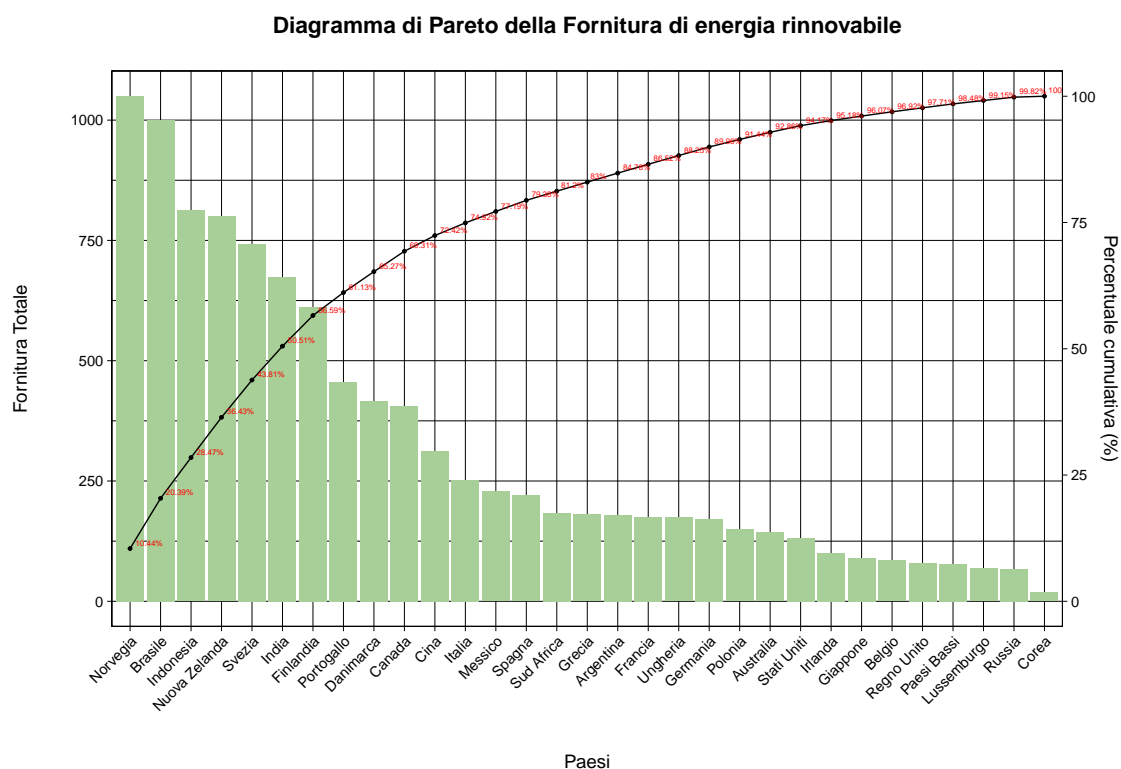


Figura 3.105: Diagramma di Pareto della variabile: Fornitura di energia rinnovabile, % del totale dell'approvvigionamento energetico

Il secondo Diagramma di Pareto, illustrato nella figura 3.106, si concentra specificamente sulla produzione di energia elettrica rinnovabile. Qui, l'analisi evidenzia un gruppo distinto di paesi che dominano il settore. **Norvegia, Brasile, Nuova Zelanda, Canada e Svezia** emergono come i principali produttori, rappresentando insieme circa il **47%** della produzione totale di energia elettrica rinnovabile tra i paesi analizzati.

Questa concentrazione evidenzia come alcuni paesi stiano effettivamente guidando il cambiamento verso una produzione di energia più verde e sostenibile. La **Norvegia**, in particolare, si distingue per la sua capacità di sfruttare risorse come l'energia idroelettrica, mentre il **Brasile** fa affidamento in larga misura su risorse come l'energia eolica e la biomassa.

Il grafico mostra anche una differenza significativa tra questi leader del settore e gli altri paesi, indicando che, nonostante i progressi complessivi, molti paesi devono ancora intraprendere passi significativi per aumentare la propria quota di energia elettrica rinnovabile.

Queste osservazioni mettono in evidenza come non tutti i paesi leader nel campo dell'energia rinnovabile siano anche all'avanguardia nella produzione di energia elettrica rinnovabile, sottolineando l'importanza di un approccio differenziato e specifico per ciascun settore delle energie rinnovabili.

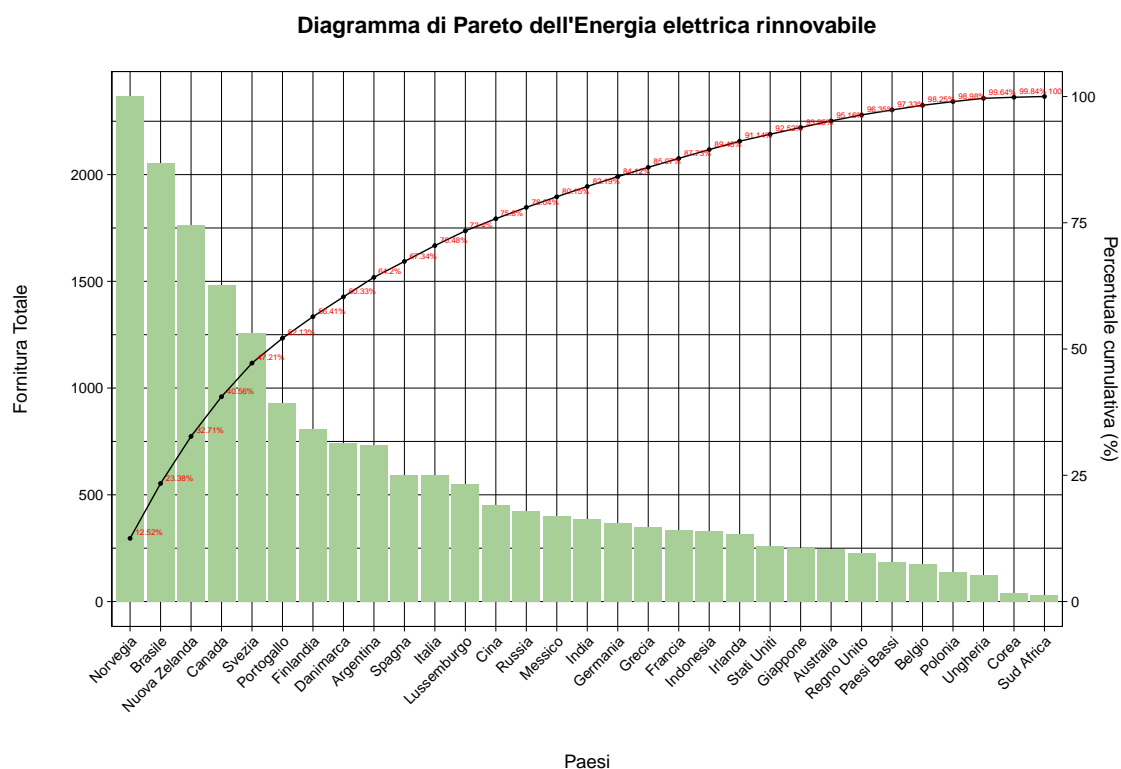


Figura 3.106: Diagramma di Pareto della variabile: Energia elettrica rinnovabile, % della generazione totale di elettricità

3.3.2 Gestione delle Risorse Idriche

Il Diagramma di Pareto relativo alla gestione delle risorse idriche fornisce un'analisi critica sulla distribuzione dello stress idrico a livello globale. La figura 3.107 evidenzia che un gruppo limitato di paesi è responsabile di una significativa parte dello stress idrico totale. In particolare, **Corea, Belgio, India, Germania, Sud Africa, Spagna e Cina** insieme costituiscono quasi il **50%** dello stress idrico tra i paesi esaminati, un dato che sottolinea la gravità della situazione in queste nazioni.

Questi paesi, caratterizzati da elevati livelli di stress idrico, si differenziano notevolmente dagli altri. Questo fenomeno può essere attribuito a vari fattori, inclusi cambiamenti climatici, crescita demografica, sviluppo industriale e agricoltura intensiva, che aumentano la pressione sulle risorse idriche disponibili.

Tuttavia, oltre a questi paesi leader, si osserva che la distribuzione dello stress idrico si equilibra gradualmente tra gli altri paesi. Ciò suggerisce che, nonostante alcune nazioni affrontino sfide idriche particolarmente gravi, lo stress idrico rimane un problema condiviso a livello globale, richiedendo interventi e strategie comuni per una gestione sostenibile delle risorse idriche.

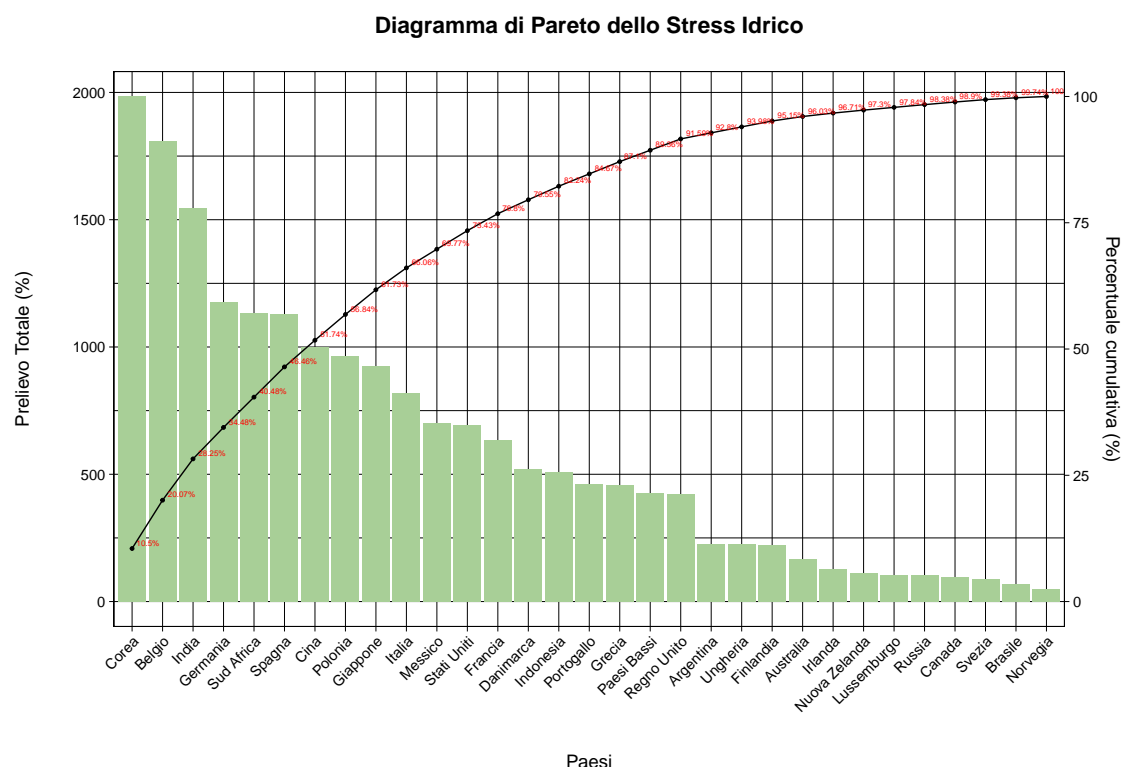


Figura 3.107: Diagramma di Pareto della variabile: Stress Idrico

3.3.3 Emissioni di CO₂

L'analisi delle emissioni di CO₂, sia basate sulla domanda sia sulla produzione, mette in luce il ruolo significativo di alcuni paesi in questo contesto ambientale globale. Come mostrato nei diagrammi di Pareto (Figura 3.108 e Figura 3.109), **Stati Uniti e Cina** emergono come i maggiori contributori alle emissioni di CO₂, rappresentando insieme circa il **50%** del totale.

In particolare, analizzando la Figura 3.108, gli **Stati Uniti** hanno totalizzato emissioni di circa **150.000 unità**, mentre la **Cina** segue a stretto giro con circa **125.000 unità**. Questi dati sottolineano la significativa disparità tra questi due paesi e gli altri, con il **Giappone**, terzo in classifica, che registra molto meno di **5.000 unità**.

L'analisi delle emissioni basate sulla produzione (Figura 3.109) mostra un ordine leggermente differente, con la **Cina** che precede gli **Stati Uniti**. Tuttavia, anche in questo scenario, l'enorme gap tra questi due paesi e gli altri rimane evidente. Oltre a **Stati Uniti e Cina**, **Russia, India, Giappone, Germania, Canada, Corea e Regno Unito** completano il gruppo dei principali emittenti, contribuendo complessivamente all'**80%** delle emissioni globali. Anche in questo caso, i valori dei paesi dal terzo in poi sono significativamente inferiori.

In entrambi i casi, i Diagrammi di Pareto mettono in evidenza la concentrazione delle emissioni di CO₂ in un numero ristretto di paesi, con un'enfasi particolare sulla predominanza di **Stati Uniti e Cina**. Questo fenomeno sottolinea la necessità di un'azione concertata e globale per affrontare le sfide climatiche, specialmente da parte dei principali contributori alle emissioni di gas serra.

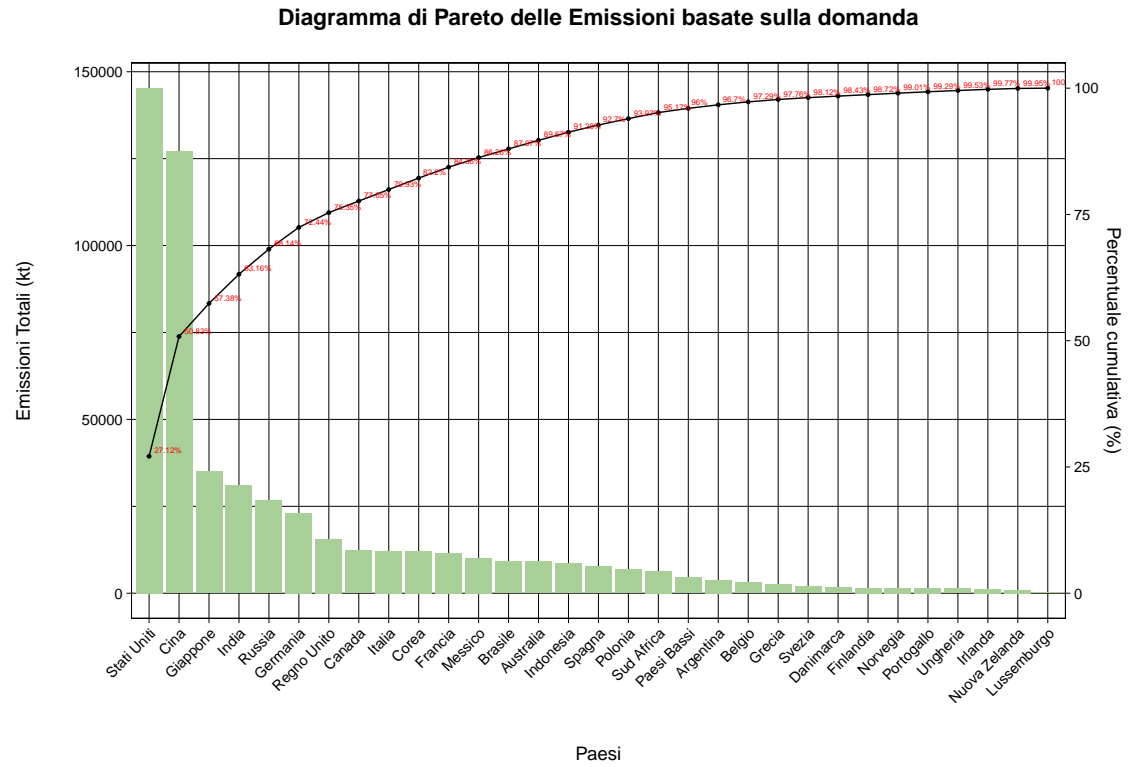


Figura 3.108: Diagramma di Pareto della variabile: Emissioni di CO₂ basate sulla domanda

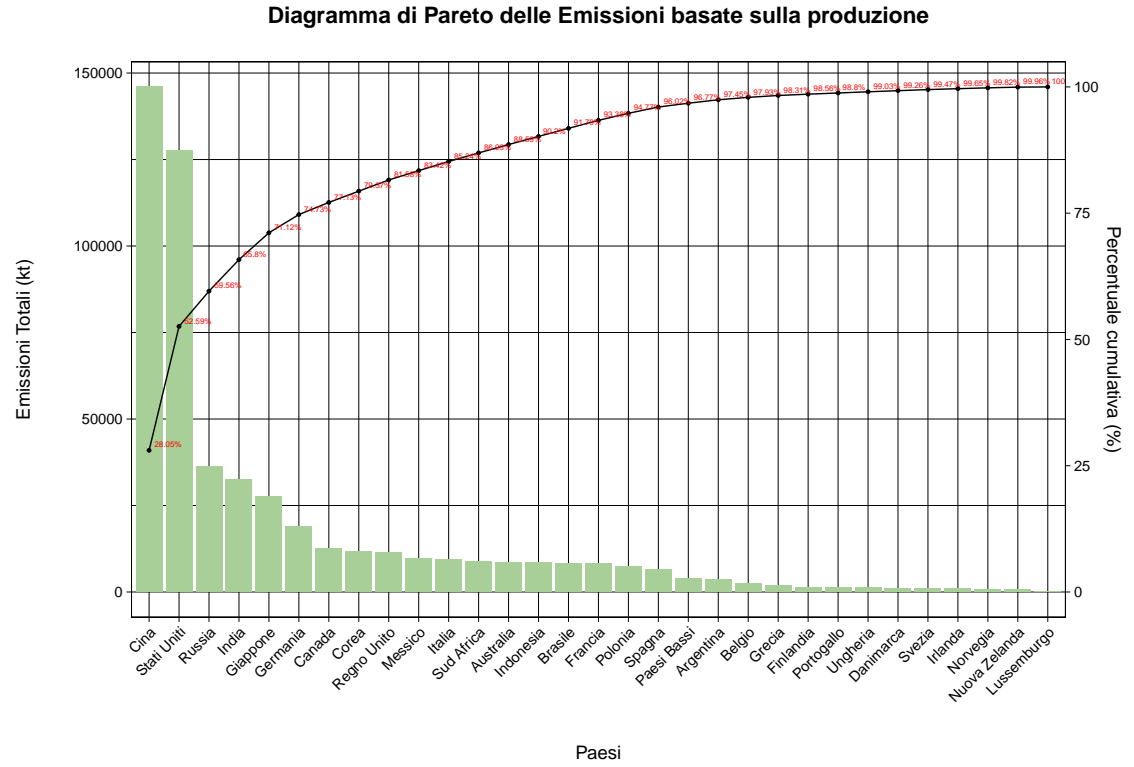


Figura 3.109: Diagramma di Pareto della variabile: Emissioni di CO₂ basate sulla produzione

3.4 Analisi Bivariata

L'**analisi bivariata** è una componente cruciale dell'analisi dei dati che si focalizza sull'esplorazione delle relazioni tra due variabili contemporaneamente. A differenza dell'analisi univariata, che esamina singolarmente ciascuna variabile, l'analisi bivariata permette di comprendere come due variabili possano essere associate tra loro. Questo tipo di analisi è fondamentale per ottenere una comprensione più profonda del comportamento dei dati, consentendo di trarre conclusioni significative che possono influenzare decisioni e strategie in una vasta gamma di settori, dalla ricerca scientifica all'ambito aziendale.

3.4.1 Matrice di Correlazione

La matrice di correlazione rappresenta uno strumento essenziale nell'analisi statistica dei dati, permettendo di esplorare in modo sistematico le relazioni tra le variabili all'interno di un dataset. Questa matrice fornisce una visione d'insieme delle correlazioni tra tutte le coppie di variabili, con ogni elemento della matrice che rappresenta il coefficiente di correlazione tra due variabili specifiche. Questo coefficiente è calcolato come il rapporto tra la covarianza campionaria delle due variabili e il prodotto delle loro deviazioni standard campionarie.

La matrice di correlazione è una matrice quadrata, in cui sia le righe che le colonne rappresentano le variabili del dataset. L'elemento situato nella riga i e nella colonna j della matrice indica la correlazione tra la variabile X_i e la variabile X_j . Il coefficiente di correlazione può variare da -1 a 1, dove valori vicini a -1 indicano una forte correlazione negativa, valori vicini a 1 indicano una forte correlazione positiva, e valori intorno a 0 suggeriscono una correlazione debole o assente.

Una delle modalità più efficaci per visualizzare una matrice di correlazione è l'utilizzo di una heatmap. Questo tipo di visualizzazione grafica consente di rappresentare i coefficienti di correlazione mediante una scala cromatica, facilitando l'identificazione immediata delle relazioni tra variabili. Nella heatmap presentata in Figura 3.110, ogni cella rappresenta la correlazione tra due variabili specifiche. Il colore della cella varia in base alla forza e alla direzione della correlazione: un colore rosso intenso denota una correlazione positiva forte, mentre un colore blu intenso indica una correlazione negativa forte. I colori intermedi rappresentano correlazioni di intensità media, mentre i colori neutri, come il bianco, indicano una correlazione debole o nulla. Questo tipo di visualizzazione è particolarmente utile per individuare rapidamente le relazioni tra le variabili del dataset.

Nella sezione seguente 3.5, verranno esaminati in dettaglio i valori di correlazione più rilevanti attraverso l'uso di scatterplot, che permettono di visualizzare in modo più dettagliato la relazione tra due variabili specifiche.

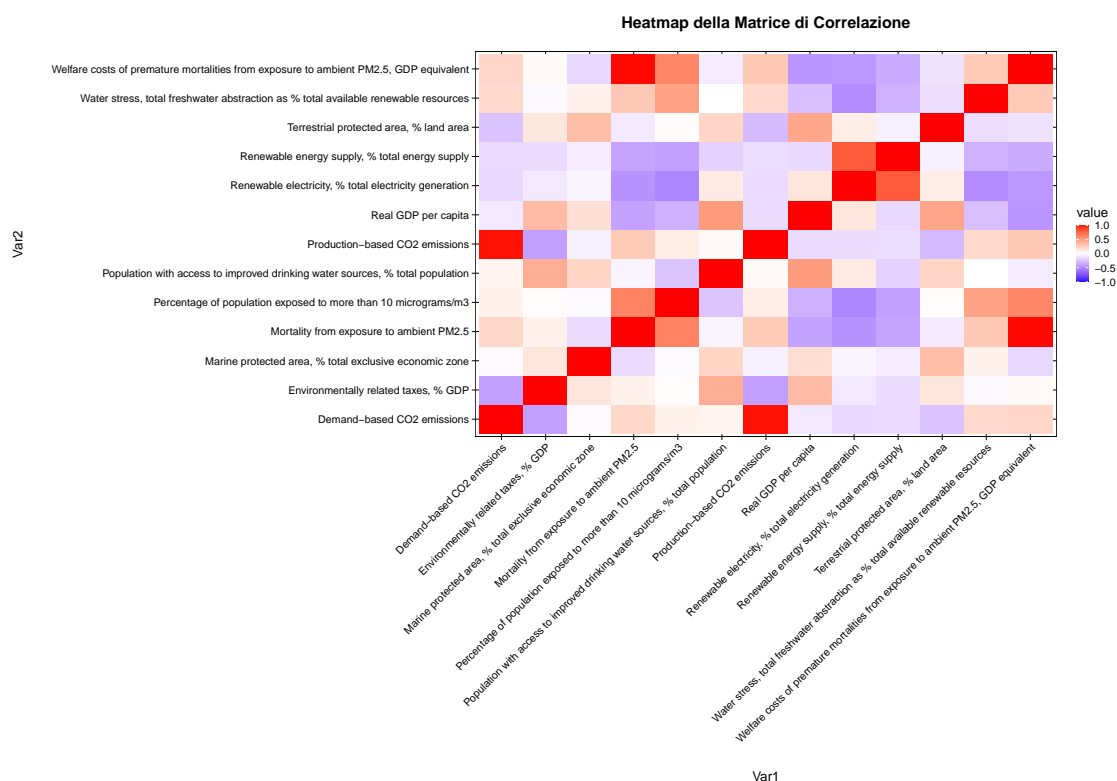


Figura 3.110: Heatmap della matrice di correlazione

3.5 Scatterplots

Uno scatterplot, o grafico a dispersione, è uno strumento visivo fondamentale nell'analisi dei dati, utilizzato per esplorare e rappresentare le relazioni tra due variabili quantitative.

In uno **scatterplot**, ogni punto sul grafico rappresenta un'osservazione individuale nel dataset. L'asse orizzontale (asse x) e l'asse verticale (asse y) rappresentano rispettivamente le due variabili che si desidera confrontare. La posizione di ciascun punto sul grafico indica i valori di queste variabili per una determinata osservazione.

3.5.1 Emissione di CO₂ basate sulla produzione vs. Emissioni di CO₂ basate sulle domande

Il grafico scatterplot 3.111 che confronta le Emissioni di CO₂ basate sulla produzione con le Emissioni di CO₂ basate sulle domande con un **coefficiente di correlazione di 0.986592** rappresenta una relazione molto forte e diretta tra queste due variabili. Un coefficiente di correlazione così elevato indica una forte correlazione lineare tra le due variabili. Questo suggerisce che, in generale, quando aumentano le emissioni di CO₂ basate sulla produzione, aumentano anche le emissioni di CO₂ basate sulle domande, e viceversa. Il fatto che queste due misurazioni delle emissioni di CO₂ siano così strettamente correlate suggerisce che per molte economie, la domanda interna è un buon indicatore delle emissioni complessive di CO₂.

La **linea di regressione** in rosso fornisce una guida visiva sulla tendenza generale dei dati e conferma la correlazione positiva tra le due variabili, confermando che un aumento nelle emissioni basate sulla produzione è associato a un aumento nelle emissioni basate sulla domanda. L'area intorno alla linea di regressione mostra l'**intervallo di errore standard** che dà un'idea della variabilità dei dati intorno alla linea di tendenza.

Ad esempio, negli **Stati Uniti**, osserviamo che le emissioni di CO₂ correlate alla produzione e alla domanda interna sono significativamente alte, riflettendo il loro ruolo di grandi consumatori e produttori di beni ad alta intensità di carbonio. Questo suggerisce che qualsiasi strategia mirata a ridurre le emissioni di CO₂ negli Stati Uniti dovrebbe considerare sia i modelli di consumo interno sia le pratiche di produzione. D'altra parte, paesi come il **Sud Africa e l'Indonesia**, pur mostrando un andamento di crescita delle emissioni simile alla tendenza generale, si discostano leggermente dalla linea di regressione. Questo potrebbe indicare una dinamica unica nella relazione tra domanda interna e produzione di CO₂, forse influenzata da fattori come la dipendenza da specifiche fonti di energia o le politiche ambientali nazionali.

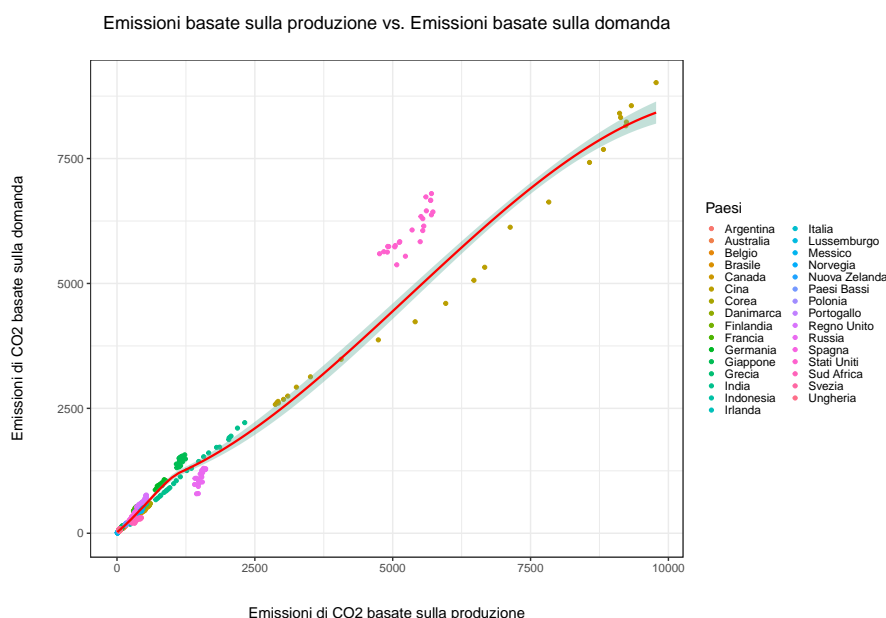


Figura 3.111: Scatterplot tra le variabili: Emissione di CO₂ basate sulla produzione vs. Emissioni di CO₂ basate sulle domande

3.5.2 Popolazione con accesso a acqua potabile migliorata e Prodotto Interno Lordo (PIL) reale pro capite

Il grafico scatterplot 3.113 che confronta la "Popolazione con accesso a acqua potabile migliorata" e il "Prodotto Interno Lordo (PIL) reale pro capite" mostra una correlazione positiva moderata, con un **coefficiente di correlazione di 0.58**. Questo suggerisce che, generalmente, le nazioni con un PIL pro capite più elevato tendono a garantire un miglior accesso all'acqua potabile alla loro popolazione. Ciò può essere interpretato come un segno che lo sviluppo economico di una nazione è spesso accompagnato da miglioramenti nelle infrastrutture essenziali, tra cui l'accesso all'acqua pulita. La **linea di regressione rossa** nel grafico mostra la tendenza generale dei dati. Il fatto che la linea sia leggermente curva potrebbe suggerire che per i paesi a più basso reddito, anche piccoli aumenti nel Prodotto Interno Lordo (PIL) pro capite possono portare a miglioramenti significativi nell'accesso all'acqua potabile, mentre per i paesi a reddito più alto, gli aumenti del Prodotto Interno Lordo (PIL) hanno un impatto minore. L'area intorno alla linea di regressione indica l'**intervallo di errore standard** e mostra dove ci si aspetterebbe che cadano la maggior parte dei punti dati. Molti paesi si trovano all'interno di questo intervallo, il che indica che seguono la tendenza generale. Tuttavia, ci sono anche diversi paesi, in particolare quelli con Prodotto Interno Lordo (PIL) pro capite più elevato, che si trovano fuori dall'intervallo e si concentrano nell'angolo in alto a destra del grafico. Questo potrebbe indicare che questi paesi hanno raggiunto un livello di accesso all'acqua quasi universale e ulteriori miglioramenti nel Prodotto Interno Lordo (PIL) pro capite non portano a grandi cambiamenti in questo indicatore.

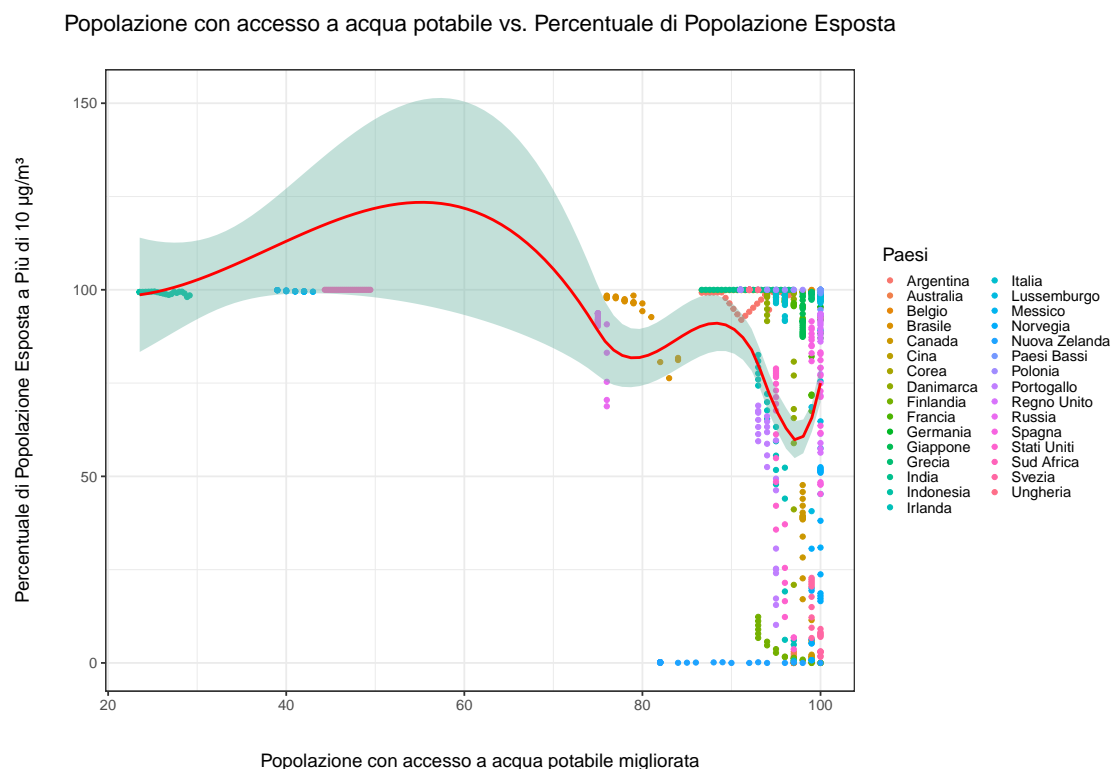


Figura 3.112: Scatterplot tra le variabili considerate: Popolazione con accesso a acqua potabile migliorata e Prodotto Interno Lordo (PIL) reale pro capite

3.5.3 Mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5) e i Costi per il benessere delle mortalità premature

Il grafico scatterplot 3.113 che mette in relazione la "Mortalità dovuta all'esposizione a Particulate Matter 2.5 (PM2.5)" con i "Costi per il benessere delle mortalità premature" mostra una correlazione estremamente forte, come evidenziato dal **coefficiente di correlazione di 0.994588**. Questo indica una relazione quasi lineare tra queste due variabili, sottolineando un legame diretto e significativo tra la salute pubblica e l'impatto economico dell'inquinamento atmosferico. La **linea di regressione rossa** nel grafico mostra che, all'aumentare dei tassi di mortalità dovuti all'esposizione a Particulate Matter 2.5 (PM2.5), aumentano proporzionalmente anche i costi per il benessere associati a queste mortalità premature. Questo suggerisce che la riduzione dei livelli di Particulate Matter 2.5 (PM2.5) potrebbe non solo migliorare la salute pubblica, ma anche ridurre i costi economici legati a queste morti premature.

In termini di politica pubblica, questo grafico fornisce un argomento per interventi mirati alla riduzione dell'inquinamento atmosferico. Non solo tali interventi possono salvare vite umane, ma possono anche contribuire a mitigare i significativi oneri economici associati alla perdita di vite e alla diminuzione della qualità della vita.

Mortalità dovuta all'esposizione a PM2.5 vs. Costi del benessere per le mortalità premature

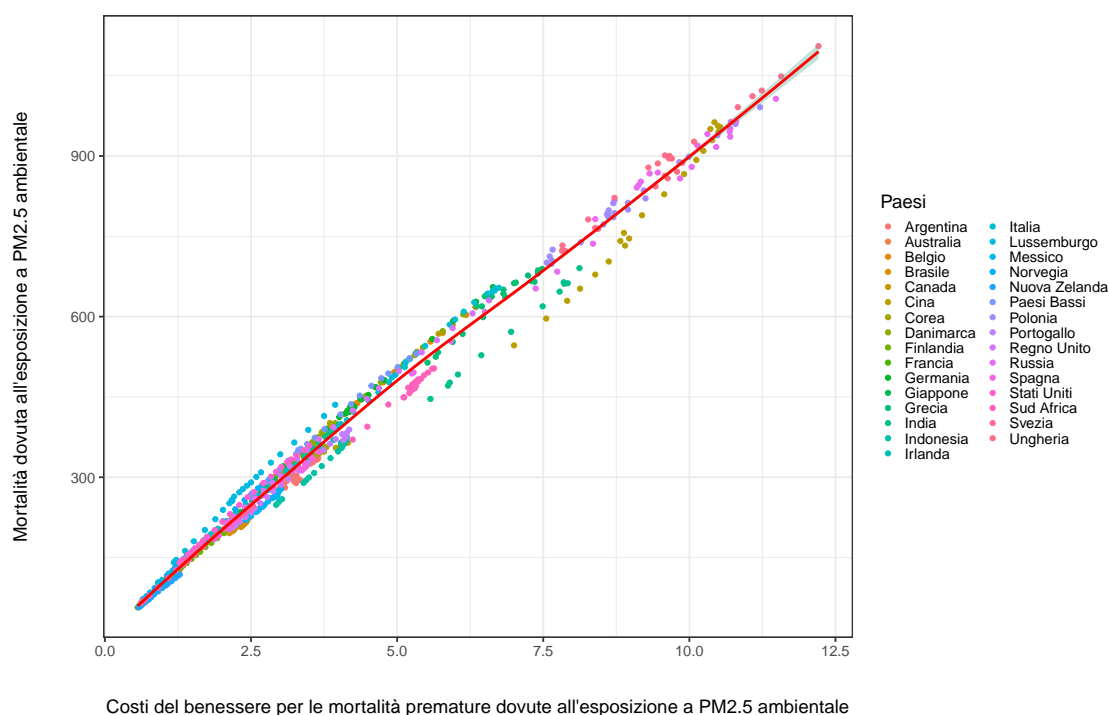


Figura 3.113: Scatterplot tra le variabili considerate

3.5.4 Energia Rinnovabile e Stress Idrico

Una delle correlazioni più significative osservate è tra Energia Rinnovabile e Stress Idrico, con un **coefficiente di correlazione di 0.631**. Questo suggerisce una forte relazione positiva, indicando che l'aumento dell'uso di energia rinnovabile potrebbe essere associato a un aumento dello stress sulle risorse idriche. Questo potrebbe essere dovuto all'uso intensivo di acqua in alcune tecnologie di energia rinnovabile, come l'idroelettrico o il raffreddamento di impianti solari termici.

La presenza di una **linea di regressione** curva nel grafico suggerisce che la relazione tra queste due variabili non è lineare, ma piuttosto che aumenti in modo più pronunciato ad alcuni livelli di consumo di energia rinnovabile. Questa curvatura potrebbe indicare che l'impatto sulle risorse idriche diventa più significativo quando il consumo di energia rinnovabile raggiunge un certo livello, possibilmente a causa dell'intensità idrica di certe forme di energia rinnovabile come l'idroelettrico e il raffreddamento dei pannelli solari termici.

L'ampio **intervallo di errore standard** mostra che, sebbene ci sia una tendenza generale, vi è una notevole variabilità nei dati. Alcuni paesi, come **l'India, il Canada e l'Australia**, mostrano valori particolarmente elevati di stress idrico associati all'uso di energia rinnovabile, deviando notevolmente dalla tendenza generale. Questo potrebbe essere dovuto a fattori locali specifici, come la dipendenza da fonti idroelettriche o la scarsità di acqua.

Inoltre, molti paesi si trovano sotto la linea di regressione, specialmente quelli con bassi livelli di stress idrico. Questo suggerisce che, per questi paesi, l'aumento dell'uso di energia rinnovabile non ha avuto un impatto significativo sullo stress idrico, forse a causa di un'ampia disponibilità di risorse idriche o dell'uso di tecnologie di energia rinnovabile meno intensive in termini di acqua.

In termini di politiche e pianificazione, questo grafico suggerisce la necessità di considerare attentamente l'equilibrio tra lo sviluppo di energia rinnovabile e la gestione sostenibile delle risorse idriche. È importante valutare attentamente i trade-off tra la riduzione delle emissioni di carbonio e l'uso sostenibile dell'acqua, specialmente in regioni con scarsità idrica o dipendenza da tecnologie idrointensive.

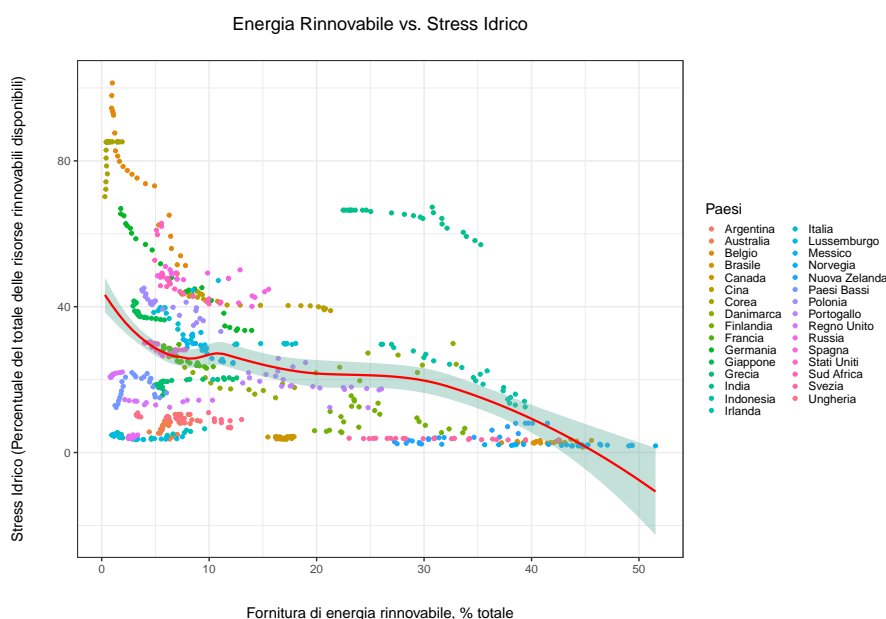


Figura 3.114: Scatterplot tra le variabili considerate

3.5.5 Energia elettrica rinnovabile, % della generazione totale di elettricità e Fornitura di energia rinnovabile, % del totale dell'approvvigionamento energetico

Con un coefficiente di correlazione di **0.7986548** tra "Energia elettrica rinnovabile, % della generazione totale di elettricità" e "Fornitura di energia rinnovabile, % del totale dell'approvvigionamento energetico" indica una forte correlazione positiva tra queste due variabili. Questo suggerisce che, nei contesti in cui una significativa percentuale dell'elettricità viene prodotta da fonti rinnovabili, vi è anche una tendenza generale verso una maggiore quota di energia rinnovabile nel mix energetico complessivo. In altre parole, i paesi che investono attivamente nella produzione di energia elettrica rinnovabile tendono anche ad avere una percentuale maggiore di energia rinnovabile nel loro approvvigionamento energetico totale. Questa forte correlazione può riflettere le strategie energetiche e le politiche ambientali che favoriscono le energie rinnovabili.

La **linea di regressione** nel grafico, leggermente curva, indica che la relazione tra le due variabili potrebbe non essere perfettamente lineare, ma piuttosto che l'incremento dell'uso di energia rinnovabile diventa più pronunciato a determinati livelli di generazione di energia elettrica rinnovabile. Questo potrebbe suggerire che, una volta superata una certa soglia di investimento o implementazione nell'energia elettrica rinnovabile, si verifica un aumento più rapido nell'utilizzo generale delle energie rinnovabili.

L'**intervallo di errore standard** mostra una certa variabilità nei dati, ma la maggior parte delle osservazioni si raggruppa intorno alla linea di regressione, indicando che la tendenza è piuttosto consistente tra i diversi paesi. Tuttavia, la presenza di osservazioni che si discostano dalla linea di regressione, in particolare quelle raggruppate vicino all'origine del grafico, indica che per alcuni paesi l'aumento nell'uso di energia elettrica rinnovabile non si è ancora tradotto in un aumento significativo della quota totale di energia rinnovabile. Questo potrebbe riflettere sfide o ritardi nell'integrazione di fonti rinnovabili nel mix energetico più ampio.

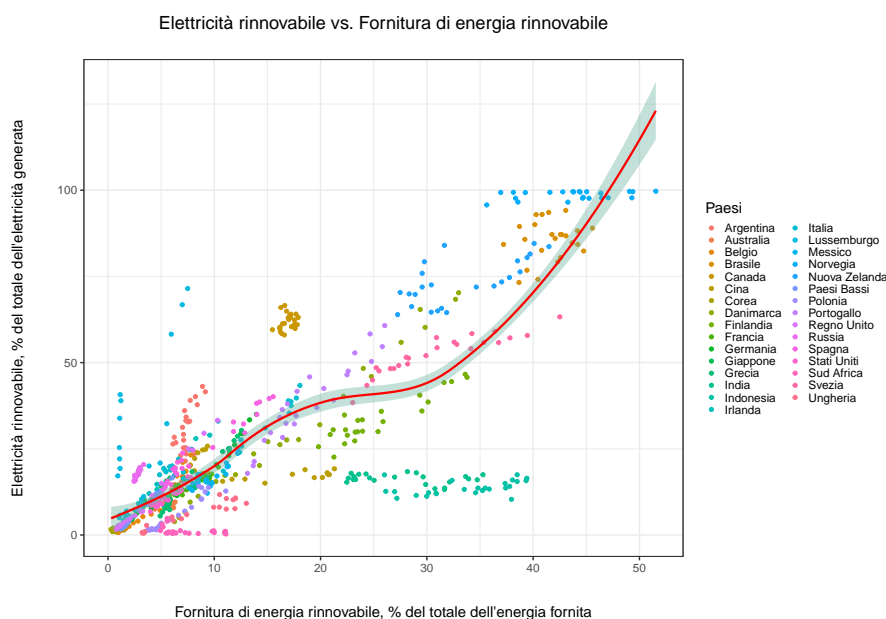


Figura 3.115: Scatterplot tra le variabili considerate

3.5.6 Percentuale della popolazione esposta a più di $10 \mu\text{g}/\text{m}^3$ e la Mortalità dovuta all'esposizione a $\text{PM}_{2.5}$ ambientali

Un coefficiente di correlazione di **0.63111061** tra la "Percentuale della popolazione esposta a più di $10 \mu\text{g}/\text{m}^3$ " e la "Mortalità dovuta all'esposizione a $\text{PM}_{2.5}$ ambientali" indica una correlazione positiva moderatamente forte. Questo valore fornisce delle intuizioni significative riguardo l'impatto dell'inquinamento atmosferico sulla salute pubblica mostra che man mano che aumenta la percentuale della popolazione esposta a livelli più alti di $\text{PM}_{2.5}$, aumenta anche il tasso di mortalità associato all'esposizione a questi inquinanti. Questa correlazione evidenzia l'importanza di interventi mirati a migliorare la qualità dell'aria. Le aree con una maggiore esposizione a $\text{PM}_{2.5}$ possono richiedere politiche più stringenti e azioni preventive per ridurre l'esposizione della popolazione e i conseguenti rischi per la salute. La correlazione può anche indicare l'importanza di considerare l'impatto dell'inquinamento atmosferico nella pianificazione urbana e nelle politiche ambientali. Ad esempio, migliorare la qualità dell'aria in aree urbane densamente popolate potrebbe avere un impatto significativo sulla riduzione della mortalità legata all'inquinamento.

La **linea di regressione** nel grafico, rimanendo abbastanza bassa, suggerisce che anche a livelli relativamente bassi di esposizione alle $\text{PM}_{2.5}$, esiste un rischio di mortalità. La distribuzione dei punti, concentrati prevalentemente sul lato destro del grafico, indica che molti paesi hanno una percentuale significativa della loro popolazione esposta a livelli di $\text{PM}_{2.5}$ superiori a $10 \mu\text{g}/\text{m}^3$. Questo è particolarmente preoccupante perché suggerisce una diffusa esposizione a un inquinante noto per i suoi effetti negativi sulla salute.

L'ampio **intervallo di errore standard** intorno alla linea di regressione indica una variabilità significativa nei dati. Ciò potrebbe essere dovuto a differenze nelle condizioni ambientali, nelle politiche di qualità dell'aria, nei livelli di sviluppo industriale e nelle pratiche di monitoraggio tra i diversi paesi. La presenza di paesi con alti livelli di esposizione e mortalità suggerisce che l'inquinamento atmosferico rimane un problema critico da affrontare, specialmente in aree urbane densamente popolate.

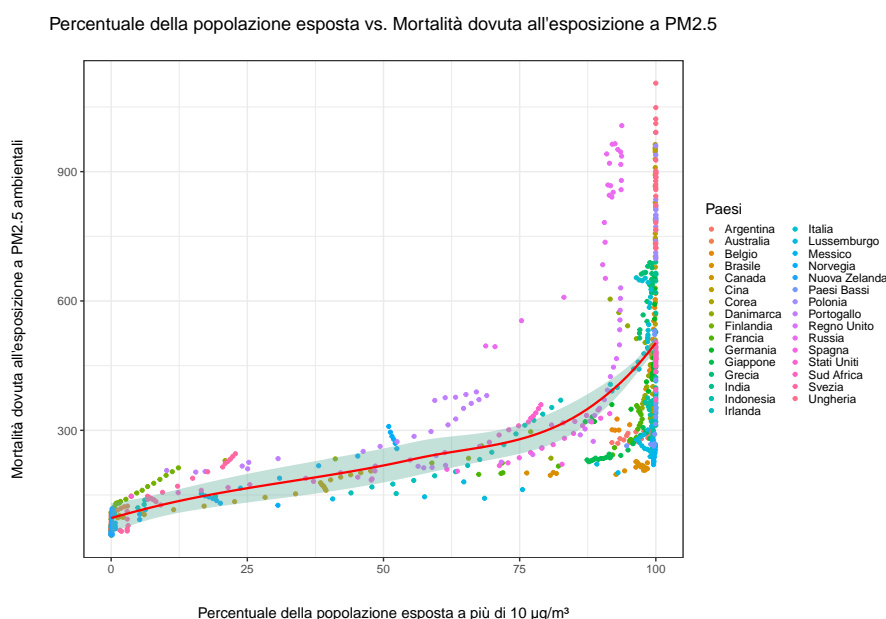


Figura 3.116: Scatterplot tra le variabili considerate

3.5.7 Costi per il benessere delle mortalità premature dovute a esposizione e la Percentuale della popolazione esposta a più di 10 $\mu\text{g}/\text{m}^3$

Un coefficiente di correlazione di **0.6114010** tra i Costi per il benessere delle mortalità premature dovute a esposizione e la Percentuale della popolazione esposta a più di 10 $\mu\text{g}/\text{m}^3$ indica una correlazione positiva moderata. Questo suggerisce che, man mano che aumenta la percentuale della popolazione esposta a un certo livello di inquinanti, aumentano anche i costi economici associati alla mortalità prematura causata da tali inquinanti. Questa correlazione evidenzia l'importante interconnessione tra la qualità dell'aria e le conseguenze economiche per la società. Indica che le aree con una maggiore esposizione a inquinanti atmosferici possono affrontare maggiori oneri economici dovuti a problemi di salute, oltre all'impatto diretto sulla salute dei cittadini. La **linea di regressione** nel grafico sale fino a metà, indicando che i costi per il benessere aumentano man mano che cresce la percentuale della popolazione esposta a inquinanti. Il fatto che la linea di regressione si stabilizzi nella seconda metà del grafico potrebbe suggerire che, oltre un certo punto, l'aumento dell'esposizione non si traduce in un incremento proporzionale dei costi per il benessere.

La presenza di paesi con alti costi per il benessere e alta esposizione suggerisce che le politiche di qualità dell'aria devono essere urgentemente intensificate in queste regioni. Le strategie potrebbero includere la promozione di trasporti pubblici più puliti, lo sviluppo di infrastrutture verdi urbane e la regolamentazione più rigorosa delle emissioni industriali.

È importante notare che l'associazione osservata non implica necessariamente una causalità diretta. Altri fattori, come le condizioni socioeconomiche e la disponibilità di infrastrutture sanitarie, possono svolgere un ruolo significativo nell'influenzare sia l'esposizione all'inquinamento che i costi associati alla salute.

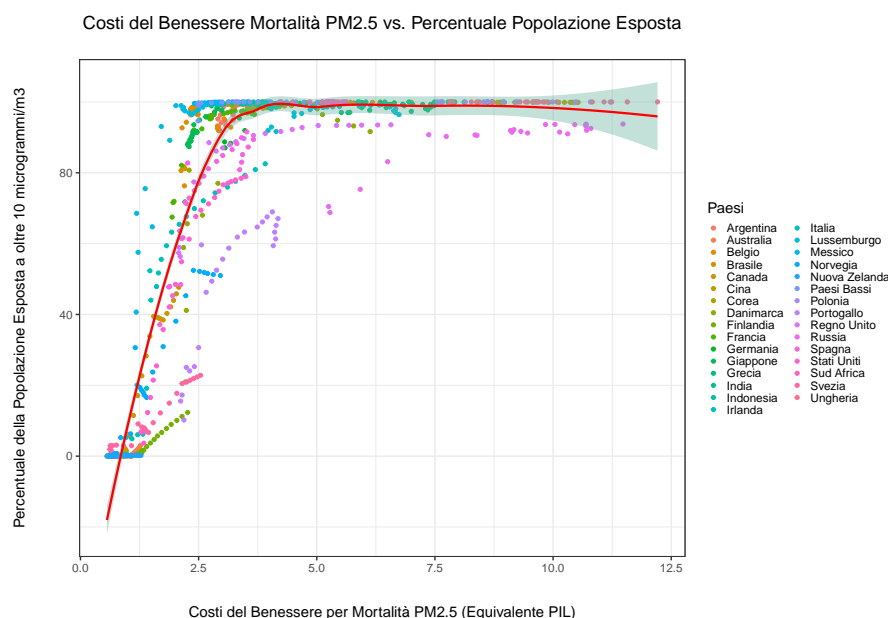


Figura 3.117: Scatterplot tra le variabili considerate

3.5.8 Emissioni di CO₂ Basate sulla Domanda vs. Tasse Correlate all'Ambiente (% del Prodotto Interno Lordo (PIL))

L'osservazione di una correlazione negativa di -0.40932837 tra le "Emissioni di CO₂ Basate sulla Domanda" e le "Tasse Correlate all'Ambiente (% del Prodotto Interno Lordo (PIL))" suggerisce che potrebbe esserci una relazione inversa tra questi due fattori. Questo significa che, in alcuni casi, un aumento delle tasse ambientali potrebbe essere associato a una diminuzione delle emissioni di CO₂ basate sulla domanda.

La distribuzione delle osservazioni prevalentemente a destra del grafico potrebbe indicare che molti paesi hanno emissioni di CO₂ basate sulla domanda relativamente elevate, ma variazioni meno pronunciate nelle tasse ambientali relative al Prodotto Interno Lordo (PIL). Questo può riflettere l'efficacia variabile delle politiche fiscali ambientali tra diversi paesi e le differenze nel loro impatto sulle emissioni di CO₂.

La linea di regressione, collocata al centro con un ampio intervallo di errore, suggerisce una relazione non molto forte tra le due variabili. Inoltre, la presenza di poche osservazioni lungo la linea di regressione indica che la correlazione potrebbe non essere uniformemente applicabile a tutti i paesi.

In sintesi, questa correlazione suggerisce che, mentre le politiche fiscali ambientali possono avere un impatto sulle emissioni di CO₂, l'effetto non è uniformemente osservabile in tutti i contesti nazionali. Potrebbe essere influenzato da una serie di fattori, inclusi il livello di sviluppo economico, le politiche ambientali specifiche, le fonti di energia predominanti e la struttura industriale. Questo grafico evidenzia l'importanza di considerare il contesto locale e le specificità di ogni paese nella progettazione di politiche fiscali per la riduzione delle emissioni di CO₂.

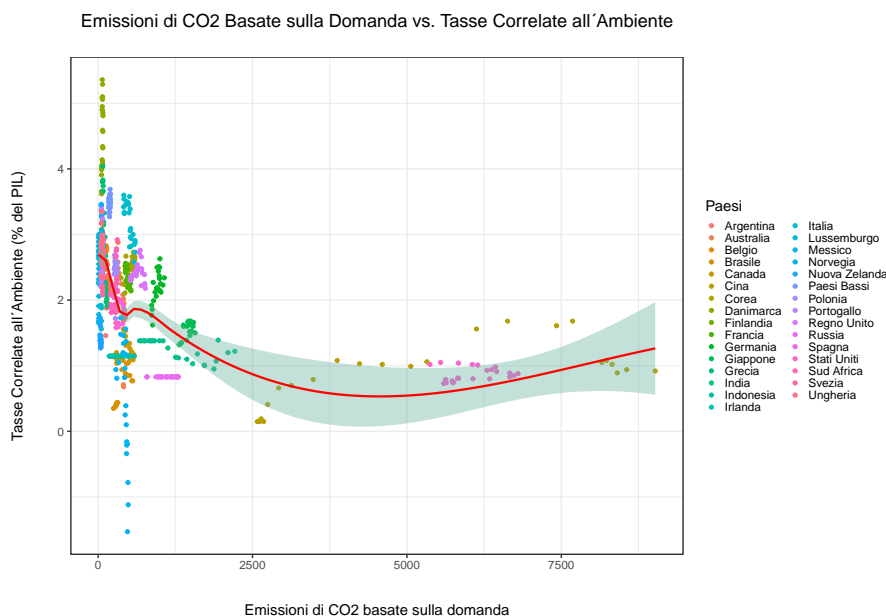


Figura 3.118: Scatterplot tra le variabili considerate

3.5.9 Prodotto Interno Lordo (PIL) Reale Pro Capite vs. Mortalità da Esposizione a Particulate Matter 2.5 (PM2.5) Ambientali

La correlazione negativa di -0.39994168 tra il PIL Reale Pro Capite e la Mortalità da Esposizione a PM2.5 Ambientali suggerisce una relazione inversa tra lo sviluppo economico di un paese e i tassi di mortalità causati dall'inquinamento atmosferico. Questo potrebbe significare che, in generale, i paesi con un PIL pro capite più elevato tendono a registrare tassi di mortalità più bassi dovuti all'esposizione a PM2.5.

La **linea di regressione** discendente da sinistra verso destra nel grafico conferma questa tendenza. La presenza di molte osservazioni sia sopra che sotto la linea di regressione indica una certa variabilità nel modo in cui differenti paesi rispondono a questa correlazione. Questo potrebbe riflettere il fatto che, oltre al livello di sviluppo economico, ci sono altri fattori importanti che influenzano la mortalità da inquinamento atmosferico, come le politiche ambientali, l'urbanizzazione, le fonti di energia e la qualità dell'aria.

In particolare, molte osservazioni relative al **Lussemburgo** in alto a destra del grafico potrebbe indicare che questo paese ha sia un alto PIL pro capite sia tassi di mortalità relativamente bassi legati all'inquinamento atmosferico, suggerendo una gestione efficace dei problemi ambientali o una migliore qualità dell'aria complessiva. L'ampio **intervallo di errore** osservato nella parte finale della linea di regressione suggerisce che la relazione tra PIL pro capite e mortalità da PM2.5 può diventare meno prevedibile a livelli di reddito più elevati. Potrebbe esserci un punto oltre il quale ulteriori aumenti del PIL pro capite non corrispondono a significative riduzioni della mortalità da inquinamento atmosferico, indicando la necessità di strategie mirate per combattere l'inquinamento oltre al semplice sviluppo economico.

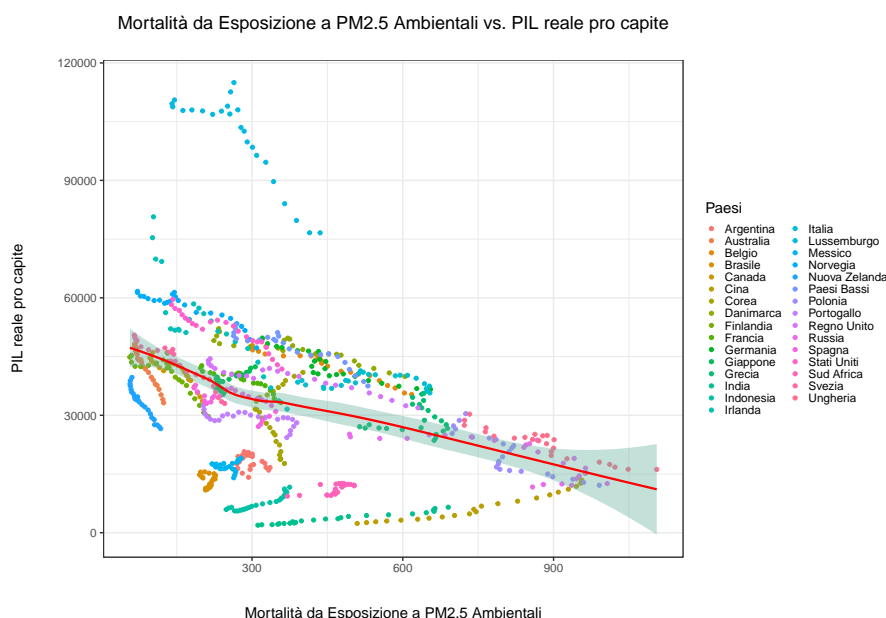


Figura 3.119: Scatterplot tra le variabili considerate

3.5.10 Popolazione con Accesso a Acqua Potabile Migliorata (% della Popolazione Totale) vs. Percentuale di Popolazione Esposta a Più di 10 $\mu\text{g}/\text{m}^3$:

La **correlazione negativa di -0.28122363** tra la Popolazione con Accesso a Acqua Potabile Migliorata e la Percentuale di Popolazione Esposta a Più di 10 $\mu\text{g}/\text{m}^3$ suggerisce una relazione inversa tra questi due indicatori di benessere e salute ambientale. Questo implica che, in generale, i paesi che riescono a fornire un maggiore accesso a fonti di acqua potabile migliorata tendono anche ad avere una minor percentuale di popolazione esposta a livelli elevati di PM2.5.

La **linea di regressione** che si presenta alta all'inizio e poi scende verso la fine del grafico, creando una curva, potrebbe indicare che questa relazione è più pronunciata in certi contesti o a determinati livelli di sviluppo. Ad esempio, potrebbe esserci un punto in cui ulteriori miglioramenti nell'accesso all'acqua potabile non corrispondono a significative riduzioni nell'esposizione agli inquinanti atmosferici, forse a causa di altri fattori ambientali o infrastrutturali.

Il fatto che molte osservazioni si trovino lungo il bordo sinistro del grafico potrebbe riflettere la presenza di numerosi paesi con accesso quasi universale a acqua potabile migliorata, ma con vari livelli di esposizione a inquinanti atmosferici. Questo potrebbe indicare che, pur avendo ottenuto grandi successi nell'accesso all'acqua potabile, questi paesi possono ancora affrontare sfide significative nella gestione della qualità dell'aria.

L'**ampio intervallo di errore** suggerisce una notevole variabilità tra i paesi nella relazione tra questi due indicatori. Questo potrebbe significare che altri fattori, come la densità della popolazione, le politiche ambientali, il tipo di industrializzazione e le fonti di inquinamento, giocano un ruolo cruciale nel determinare il livello di esposizione agli inquinanti atmosferici, indipendentemente dal successo nell'accesso all'acqua potabile.

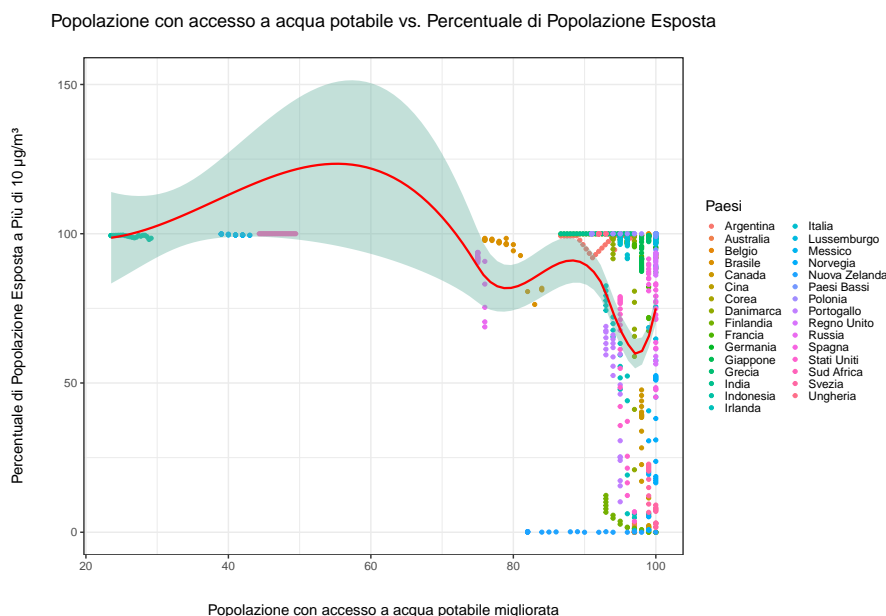


Figura 3.120: Scatterplot tra le variabili considerate

3.5.11 Energia Elettrica Rinnovabile (Percentuale della Generazione Totale di Elettricità) vs. Percentuale di Popolazione Esposta a Più di 10 $\mu\text{g}/\text{m}^3$

La **correlazione negativa di -0.52928322** tra l'Energia Elettrica Rinnovabile (Percentuale della Generazione Totale di Elettricità) e la Percentuale di Popolazione Esposta a Più di 10 $\mu\text{g}/\text{m}^3$ indica una relazione inversa tra l'uso di energia elettrica rinnovabile e l'esposizione della popolazione a livelli più elevati di PM2.5.

La **linea di regressione curva** che scende da sinistra a destra potrebbe suggerire che questa relazione non è lineare, ma varia in modo più significativo in alcune fasce di utilizzo di energia elettrica rinnovabile. Ad esempio, potrebbe esserci un punto in cui ulteriori aumenti nell'uso di energia elettrica rinnovabile portano a significative riduzioni nell'esposizione ai PM2.5, ma questa relazione potrebbe indebolirsi a livelli molto alti di utilizzo di energia rinnovabile.

La presenza di molte osservazioni lungo il bordo sinistro del grafico potrebbe indicare che esistono molti paesi con una bassa percentuale di popolazione esposta a livelli elevati di PM2.5, indipendentemente dall'uso di energia elettrica rinnovabile. Questo suggerisce che altri fattori, come la politica ambientale, la geografia e la densità di popolazione, giocano un ruolo nel determinare il livello di esposizione agli inquinanti atmosferici.

In generale, questo grafico suggerisce che c'è una tendenza che collega positivamente l'uso di energia elettrica rinnovabile con una minore esposizione alla contaminazione atmosferica, ma anche che questa relazione è complessa e influenzata da una serie di fattori. Rafforzare gli investimenti nell'energia elettrica rinnovabile potrebbe contribuire a migliorare la qualità dell'aria e la salute pubblica, ma è necessario considerare l'interazione di vari elementi ambientali e politiche per ottenere risultati significativi.

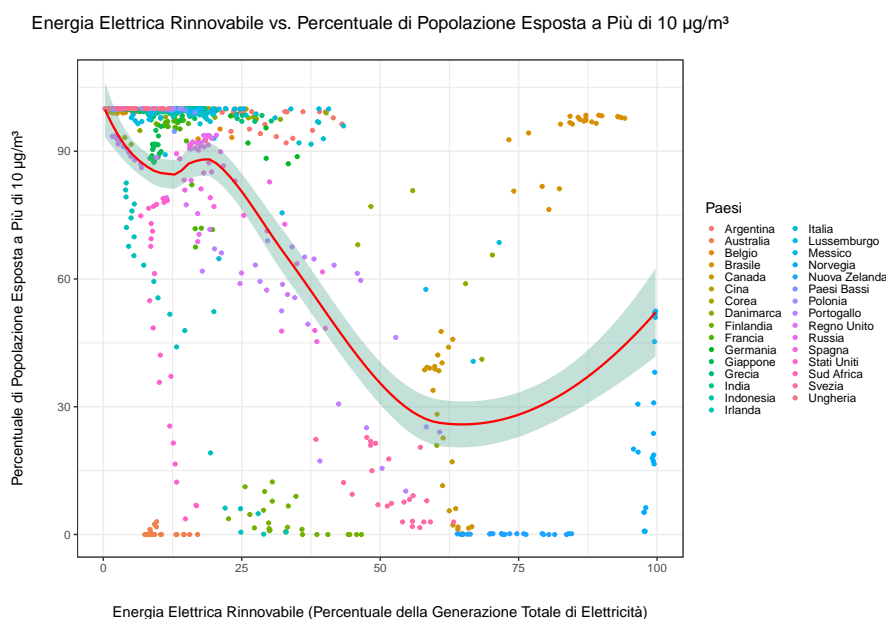


Figura 3.121: Scatterplot tra le variabili considerate

Capitolo 4

Regressione

La modellazione tramite regressione rappresenta uno strumento statistico fondamentale per analizzare e comprendere le relazioni quantitative tra variabili in contesti complessi, come l'inquinamento atmosferico e le sue conseguenze sulla salute pubblica. In questi ambiti, dove numerosi fattori interagiscono tra loro, la capacità di identificare e quantificare le connessioni tra variabili è cruciale per sviluppare politiche e interventi efficaci.

In questo capitolo, ci proponiamo di esplorare le dinamiche che legano diverse variabili chiave, quali:

- **Emissioni di CO₂:** Le emissioni di anidride carbonica rappresentano uno dei principali fattori contribuenti al cambiamento climatico. Analizzeremo le relazioni tra le emissioni basate sulla produzione e quelle basate sulla domanda, al fine di comprendere come le attività economiche e produttive influenzano la domanda globale di CO₂.
- **Produzione di Energia Rinnovabile:** L'energia rinnovabile è considerata una delle soluzioni chiave per mitigare l'inquinamento atmosferico e ridurre la dipendenza dai combustibili fossili. Esamineremo come la crescita nella produzione di energia rinnovabile influisce sulla riduzione delle emissioni inquinanti e sull'esposizione della popolazione a PM2.5.
- **Esposizione della Popolazione a Inquinanti Atmosferici:** La qualità dell'aria ha un impatto diretto sulla salute pubblica. Utilizzando modelli di regressione, studieremo la relazione tra la produzione di energia, l'esposizione a PM2.5 e le conseguenze sanitarie, come la mortalità prematura.
- **Impatto Economico e Sanitario dell'Inquinamento:** Oltre ai costi umani, l'inquinamento atmosferico comporta significativi costi economici. Analizzeremo la relazione tra i costi del benessere associati alla mortalità prematura e le variabili ambientali, fornendo stime utili per valutare l'efficacia delle politiche di intervento.

4.1 Emissioni Prodotte vs. Domandate

In questa sezione, esaminiamo la relazione tra le emissioni di CO₂ basate sulla produzione e quelle basate sulla domanda utilizzando un modello di regressione lineare semplice. I risultati di questa analisi ci permettono di comprendere come la produzione di CO₂ influenzi la domanda di emissioni, un'informazione cruciale per sviluppare politiche ambientali efficaci.

4.1.1 Risultati del Modello

Dall'analisi del modello di regressione lineare semplice emerge che la relazione tra le emissioni di CO₂ basate sulla produzione (variabile indipendente) e quelle basate sulla domanda (variabile dipendente) può essere espressa dalla seguente equazione:

$$\text{Emissioni basate sulla domanda} = 42.03 + 0.97 \times \text{Emissioni basate sulla produzione}$$

Dove:

- 42.03 è l'**intercetta del modello**, che rappresenta il valore delle emissioni basate sulla domanda quando le emissioni basate sulla produzione sono zero. L'intervallo di confidenza al 95% per questo valore è compreso tra 23.42 e 60.63. Il p-value associato all'intercetta è 1.07×10^{-5} , che è molto inferiore al livello di significatività comunemente utilizzato (ad esempio, 0.05). Questo indica che l'intercetta è altamente significativa e che la relazione tra la variabile indipendente (emissioni di CO₂ basate sulla produzione) e la variabile dipendente (emissioni di CO₂ basate sulla domanda) non passa attraverso l'origine del grafico, ma piuttosto si stabilizza a un livello di base di 42.03 unità di CO₂ basate sulla domanda, anche quando le emissioni di produzione sono pari a zero.
- 0.97 è il **coefficiente di regressione**, che indica che per ogni unità di aumento delle emissioni basate sulla produzione, le emissioni basate sulla domanda aumentano di 0.97 unità. L'intervallo di confidenza al 95% per questo coefficiente è compreso tra 0.96 e 0.98. Il p-value per il coefficiente della produzione di CO₂ è estremamente basso, inferiore a 2×10^{-16} . Questo indica una fortissima evidenza statistica a favore dell'esistenza di una relazione lineare tra la produzione di CO₂ e la domanda di CO₂.

Il modello ha un valore di R^2 pari a 0.9734, il che indica che il 97.34% della variabilità nelle emissioni basate sulla domanda può essere spiegata dalle emissioni basate sulla produzione. Questo alto valore di R^2 suggerisce una forte relazione lineare tra le due variabili.

4.1.2 Analisi dei Residui

L'errore standard dei residui del modello è pari a 232.9, indicando la deviazione media dei residui dal valore predetto dal modello.

Inoltre, l'analisi dei residui è fondamentale per valutare l'adeguatezza del modello. I residui standardizzati ci forniscono informazioni su quanto i valori previsti dal modello si discostino dai valori osservati.

Residui Standardizzati Il grafico dei residui standardizzati (Figura 4.1) mostra una **distribuzione relativamente casuale attorno allo zero**, suggerendo che il modello cattura bene la relazione tra le due variabili. Tuttavia, ci sono alcuni outlier significativi, in particolare per paesi come **Stati Uniti, Cina e Russia**, che indicano che in questi casi il modello potrebbe non adattarsi perfettamente ai dati.

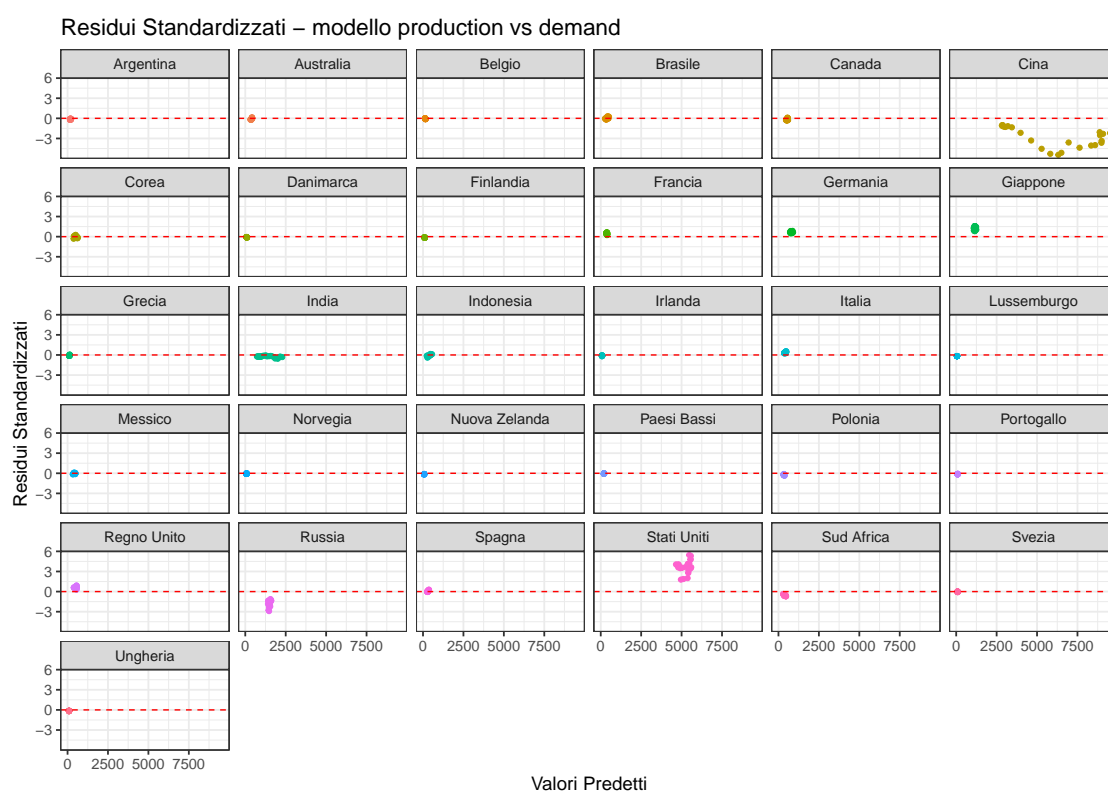


Figura 4.1: Grafico dei Residui Standardizzati per il modello tra Emissioni di CO₂ basate sulla Produzione e sulla Domanda.

QQ Plot Il QQ plot (Figura 4.2) consente di valutare la normalità dei residui. La maggior parte dei punti segue la linea di riferimento, indicando che i residui si **distribuiscono in modo approssimativamente normale**, un'ulteriore conferma della bontà del modello.

Scale-Location Plot Il grafico Scale-Location (Figura 4.3) mostra la radice quadrata dei residui assoluti rispetto ai valori predetti. Questo grafico evidenzia che non ci sono particolari problemi di eteroschedasticità (varianza non costante dei residui), anche se alcuni paesi mostrano una leggera deviazione.

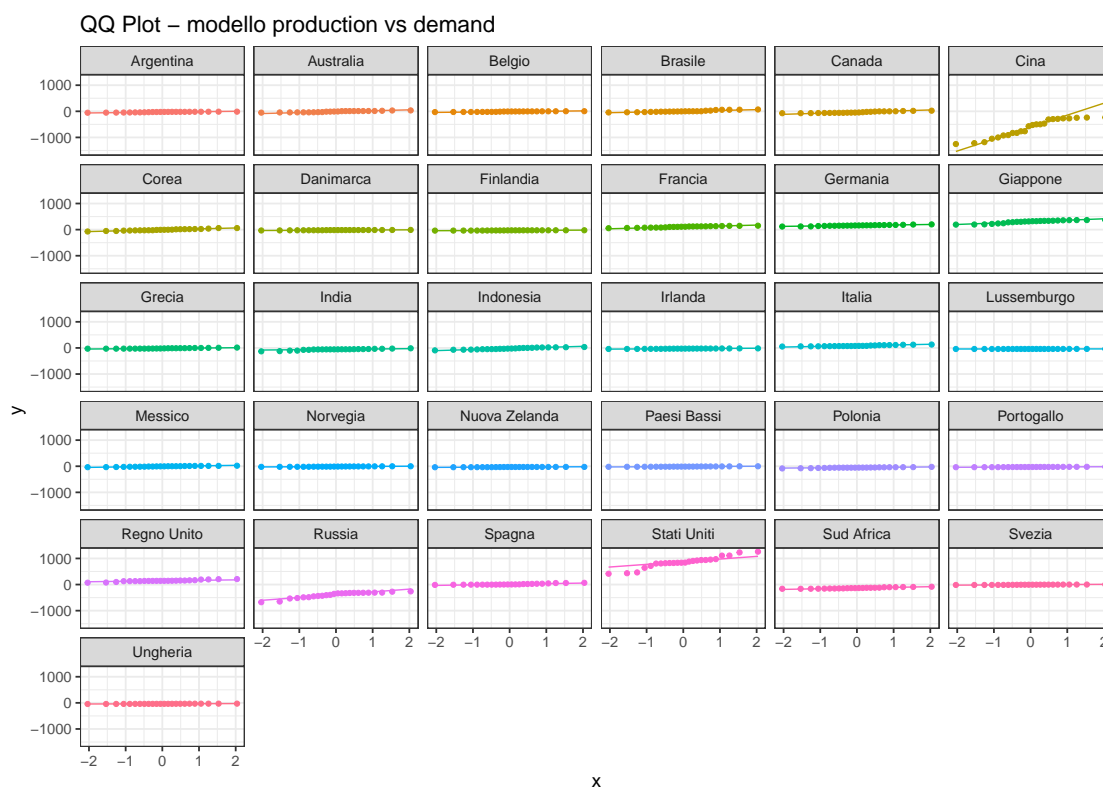


Figura 4.2: QQ Plot dei residui per il modello tra Emissioni CO₂ di basate sulla Produzione e sulla Domanda.

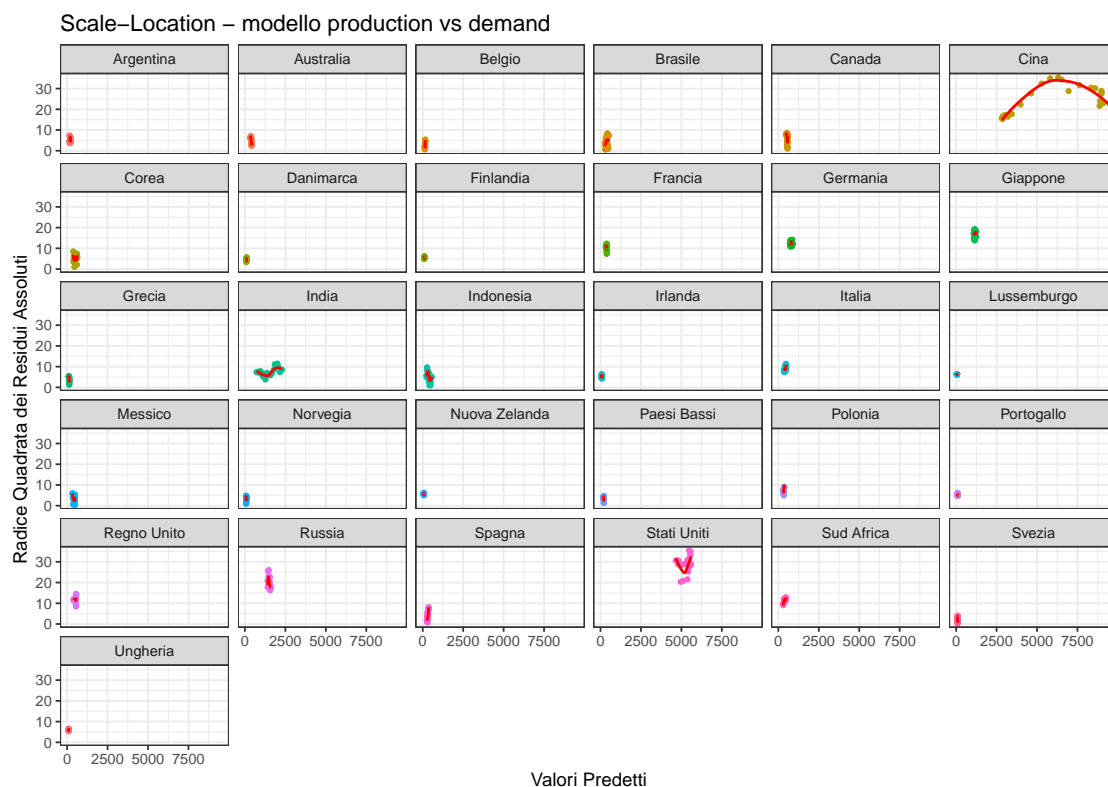


Figura 4.3: Grafico Scale-Location per il modello tra Emissioni di CO₂ basate sulla Produzione e sulla Domanda.

Il modello di regressione lineare tra le emissioni di CO₂ basate sulla produzione e sulla domanda si dimostra altamente significativo, con un alto valore di R^2 che suggerisce una forte relazione lineare. Tuttavia, la presenza di outlier significativi in alcuni paesi suggerisce la necessità di ulteriori analisi, magari includendo variabili aggiuntive o esplorando modelli non lineari per migliorare l'adattamento del modello.

Questi risultati sono fondamentali per la definizione di politiche ambientali efficaci, poiché indicano che la riduzione delle emissioni di CO₂ nella produzione potrebbe avere un impatto diretto e sostanziale sulla domanda complessiva di emissioni.

4.2 Accesso all'Acqua Potabile e PM2.5

In questa sezione, esamineremo la relazione tra la **percentuale di popolazione con accesso a fonti di acqua potabile migliorate** e la **percentuale di popolazione esposta a più di 10 µg/m³ di PM2.5** utilizzando un modello di regressione polinomiale di secondo grado. L'obiettivo è comprendere in che modo la disponibilità di acqua potabile migliorata possa influenzare l'esposizione della popolazione a livelli pericolosi di PM2.5, un inquinante atmosferico noto per i suoi gravi effetti sulla salute.

Il modello di regressione polinomiale di secondo grado è stato costruito come segue:

$$\text{Esposizione PM2.5(\%)} = \alpha + \beta_1 \times (\text{Acqua Potabile Migliorata (\%)}) + \beta_2 \times (\text{Acqua Potabile Migliorata (\%)})^2$$

Dove:

- α rappresenta l'**intercetta del modello**, ovvero la stima della percentuale di popolazione esposta a PM2.5 quando l'accesso a fonti di acqua potabile migliorate è pari a zero.
- β_1 è il **coefficiente associato al termine lineare**, che misura l'impatto diretto dell'incremento di un'unità percentuale nell'accesso a fonti di acqua potabile migliorate sulla percentuale di popolazione esposta a PM2.5.
- β_2 è il **coefficiente associato al termine quadratico**, che cattura eventuali effetti non lineari dell'accesso a fonti di acqua potabile migliorate sull'esposizione a PM2.5.

4.2.1 Risultati del Modello

I risultati ottenuti dall'analisi del modello di regressione polinomiale sono i seguenti:

- **Intercetta (α):** Il valore stimato dell'intercetta è 74.204, con un errore standard di 1.295. Questo valore è altamente significativo, come indicato dal p-value inferiore a 2×10^{-16} , suggerendo che esiste una percentuale sostanziale di popolazione esposta a PM2.5 anche quando l'accesso all'acqua potabile migliorata è minimo o nullo.
- **Coefficiente del termine lineare (β_1):** Il coefficiente stimato per il termine lineare è -245.207, con un errore standard di 35.323. Questo coefficiente è anch'esso altamente significativo (p-value 8.46×10^{-12}), indicando che un incremento nella percentuale di popolazione con accesso a fonti di acqua potabile migliorate è associato a una significativa riduzione nella percentuale di popolazione esposta a livelli pericolosi di PM2.5.

- **Coefficiente del termine quadratico (β_2):** Il coefficiente stimato per il termine quadratico è -40.053 , con un errore standard di 35.323 . Tuttavia, questo coefficiente non risulta significativo (p-value 0.257), suggerendo che l'inclusione del termine quadratico non aggiunge un contributo significativo alla spiegazione della varianza nell'esposizione aPM2.5.
- **Errore standard dei residui:** L'errore standard dei residui è pari a 35.32 , il che indica la dispersione media dei residui attorno ai valori predetti dal modello.
- **R-quadrato:** Il valore di R^2 è 0.0626 , suggerendo che il modello spiega solo circa il 6.26% della variabilità nella percentuale di popolazione esposta aPM2.5. Questo basso valore di R^2 indica che esistono altri fattori significativi, non inclusi nel modello, che influenzano l'esposizione della popolazione aPM2.5.

Gli intervalli di confidenza al 95% per i coefficienti del modello sono riportati di seguito:

- **Intercetta (α):** L'intervallo di confidenza è compreso tra 71.660 e 76.748 . Questo intervallo ristretto indica che la stima dell'intercetta è abbastanza precisa, confermando che il livello base di esposizione aPM2.5 è significativamente diverso da zero.
- **Primo Coefficiente Polinomiale (β_1):** L'intervallo di confidenza per il coefficiente lineare varia da -314.012 a -176.401 . Questo intervallo conferma la forte relazione inversa tra l'accesso a fonti di acqua potabile migliorate e l'esposizione aPM2.5, con una stima del coefficiente significativamente diversa da zero.
- **Secondo Coefficiente Polinomiale (β_2):** L'intervallo di confidenza per il coefficiente quadratico è compreso tra -109.237 e 29.132 . Poiché l'intervallo include lo zero, si conferma che il termine quadratico non ha un impatto significativo sulla variabile dipendente.

Il modello di regressione polinomiale suggerisce che esiste una relazione inversa significativa tra l'accesso a fonti di acqua potabile migliorate e l'esposizione della popolazione aPM2.5, almeno per quanto riguarda il termine lineare. Tuttavia, la componente quadratica non risulta significativa, indicando che una relazione quadratica non è necessaria per spiegare l'associazione tra queste variabili.

Il basso valore di R^2 segnala che gran parte della variabilità nell'esposizione a PM2.5 non è spiegata dal modello attuale, suggerendo la presenza di altri fattori determinanti non considerati.

4.2.2 Analisi dei Residui

I residui del modello sono stati analizzati attraverso vari grafici diagnostici:

Grafico dei Residui Standardizzati Come mostrato nella Figura 4.4, i residui non presentano una distribuzione perfettamente casuale, suggerendo la presenza di un pattern nei dati non completamente catturato dal modello.

QQ Plot Il QQ plot dei residui, rappresentato nella Figura 4.5, evidenzia alcune deviazioni dai quantili teorici, suggerendo che i residui non seguono perfettamente una distribuzione normale.

Grafico Scale-Location Come mostrato nella Figura 4.6, il grafico Scale-Location evidenzia una certa eterogeneità nella dispersione dei residui, con variazioni significative tra diversi paesi.

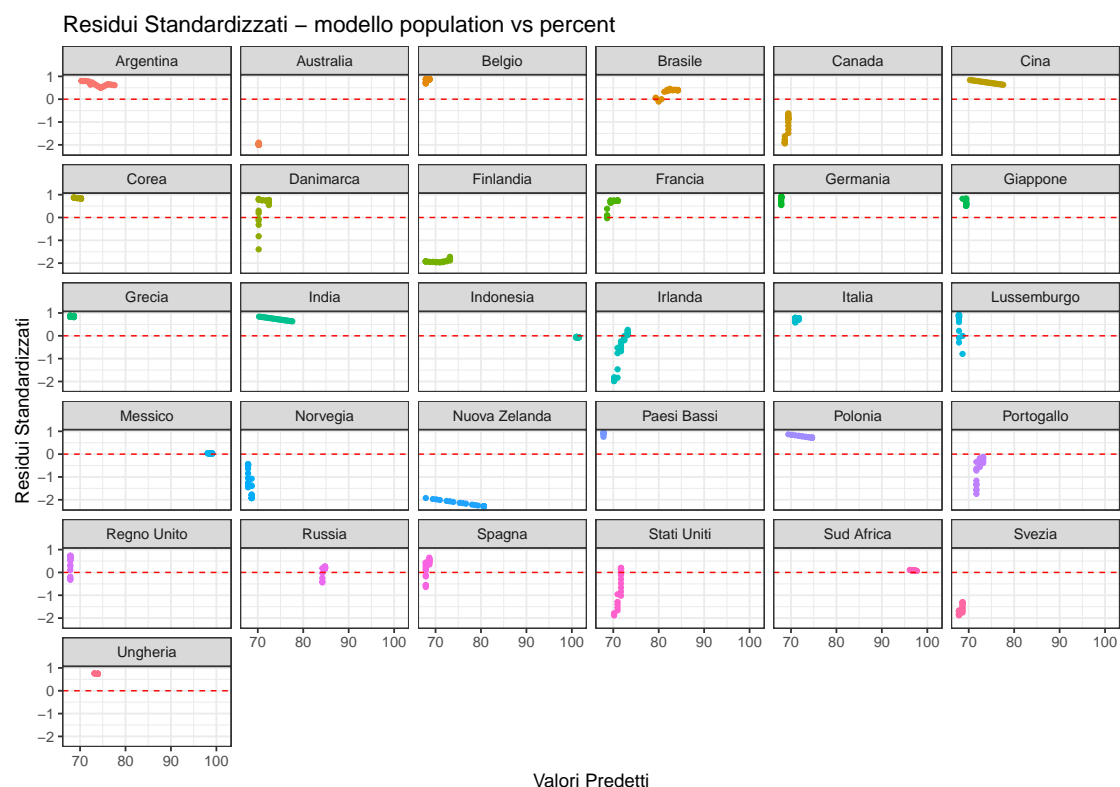


Figura 4.4: Grafico dei Residui Standardizzati per il modello Acqua Potabile Migliorata vs. Percentuale di Popolazione Esposta aPM2.5.

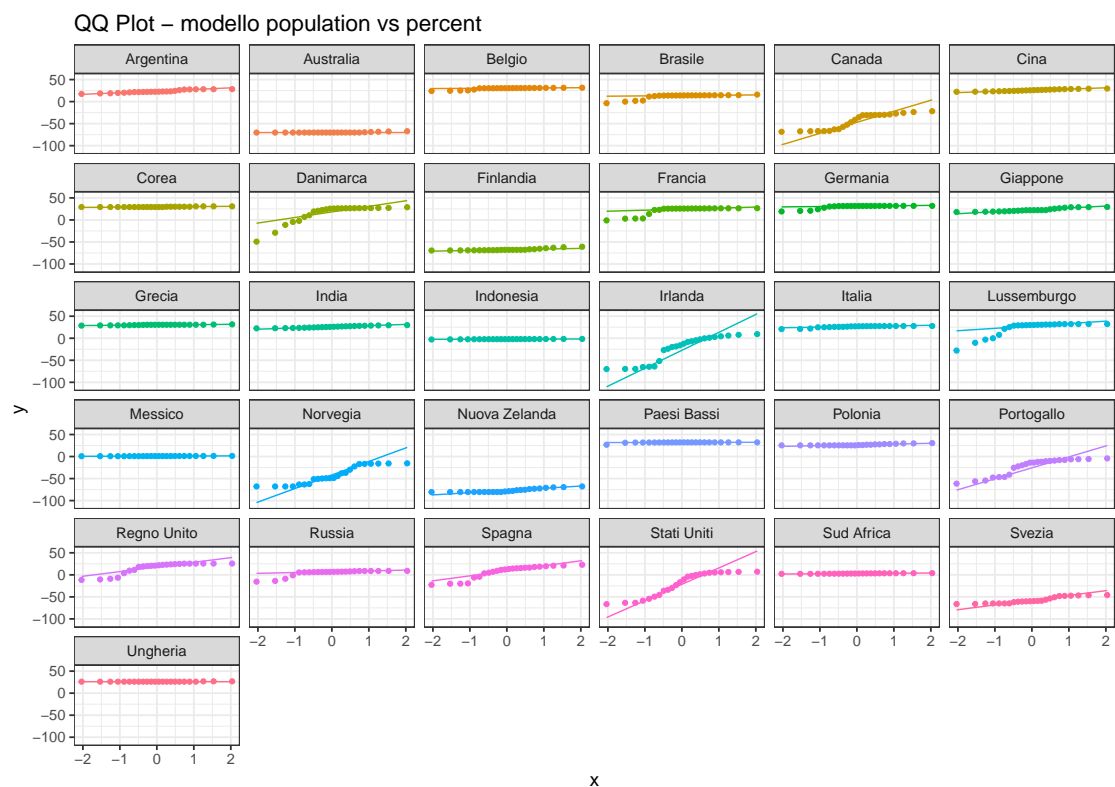


Figura 4.5: QQ Plot per il modello Acqua Potabile Migliorata vs. Percentuale di Popolazione Esposta aPM2.5.



Figura 4.6: Grafico Scale-Location per il modello Acqua Potabile Migliorata vs. Percentuale di Popolazione Esposta aPM2.5.

4.3 Elettricità Rinnovabile e PM2.5

In questa sezione, viene analizzata la relazione tra la percentuale di elettricità generata da fonti rinnovabili e la percentuale della popolazione esposta a livelli diPM2.5 superiori a 10 microgrammi per metro cubo. Il modello di regressione lineare è stato utilizzato per valutare l'impatto dell'elettricità rinnovabile sull'esposizione della popolazione a inquinanti atmosferici.

4.3.1 Risultati del Modello

Il modello di regressione lineare ha prodotto i seguenti risultati:

$$\text{Popolazione Esposta (\%)} = 93.55 - 0.76 \times \text{Elettricità Rinnovabile (\%)}$$

Dove:

- 93.55 è l'**intercetta**, che rappresenta la percentuale di popolazione esposta quando la percentuale di elettricità rinnovabile è zero.
- -0.76 è il **coefficiente per la percentuale di elettricità rinnovabile**, suggerendo che un aumento dell'1% nella percentuale di elettricità rinnovabile è associato a una riduzione dello 0.76% nella percentuale di popolazione esposta aPM2.5.

I risultati sono statisticamente significativi, come indicato dai p-value molto bassi per l'intercetta e il coefficiente ($p < 2 \times 10^{-16}$).

Il valore di R^2 di 0.28 suggerisce che circa il 28% della varianza nella percentuale di popolazione esposta è spiegata dal modello. Sebbene il modello mostri una relazione significativa tra le due variabili, una porzione significativa della varianza non è spiegata, indicando la presenza di altri fattori influenti non inclusi nel modello.

4.3.2 Analisi dei Residui

L'analisi dei residui standardizzati è stata condotta per valutare l'adattamento del modello ai dati.

Grafico dei Residui Standardizzati La Figura 4.7 mostra i residui standardizzati rispetto ai valori predetti dal modello.

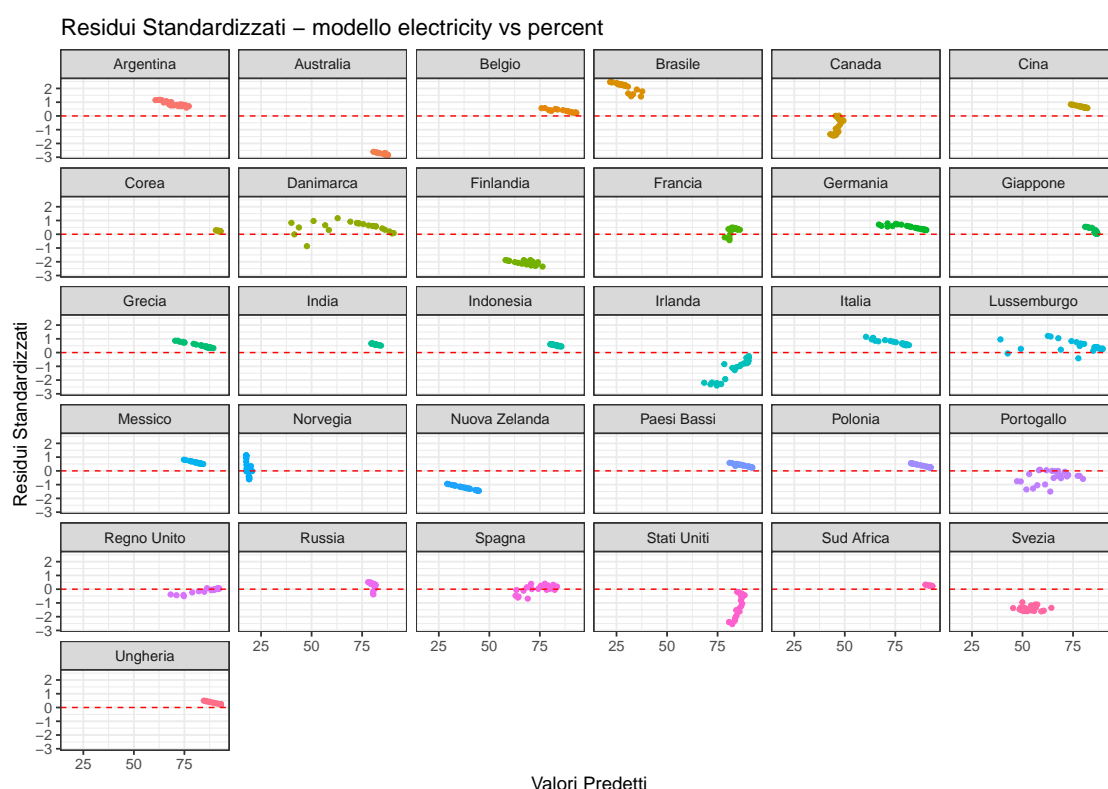


Figura 4.7: Grafico dei residui standardizzati per il modello Eletttricità Rinnovabile vs. Percentuale di Popolazione Esposta.

La distribuzione dei residui attorno allo zero suggerisce che il modello lineare cattura in modo adeguato la relazione tra elettricità rinnovabile e percentuale di popolazione esposta. Tuttavia, la presenza di alcuni outlier indica che esistono fattori non inclusi nel modello che potrebbero influenzare l'esposizione della popolazione.

QQ Plot Il QQ plot per il modello (Figura 4.8) mostra che i residui seguono una distribuzione approssimativamente normale, con qualche deviazione agli estremi.

Grafico Scale-Location Il grafico Scale-Location (Figura 4.9) è stato utilizzato per valutare l'omogeneità della varianza dei residui. L'analisi mostra che la varianza dei residui non è costante,

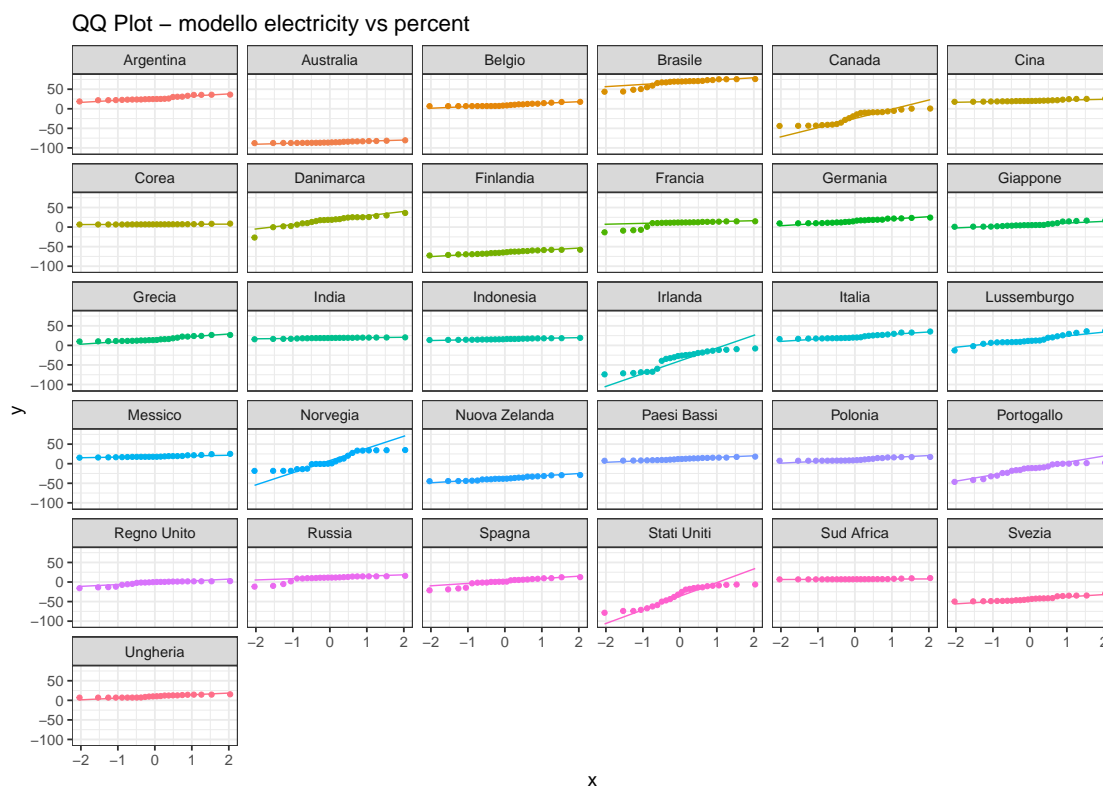


Figura 4.8: QQ Plot per il modello Elettricità Rinnovabile vs. Percentuale di Popolazione Esposta.

con alcune tendenze che suggeriscono possibili problemi di eteroschedasticità.

Questi risultati indicano che, sebbene il modello lineare fornisca una buona base per comprendere la relazione tra elettricità rinnovabile e esposizione della popolazione aPM2.5, ulteriori ricerche potrebbero essere necessarie per affrontare i problemi di varianza non costante e migliorare l'adattamento del modello.

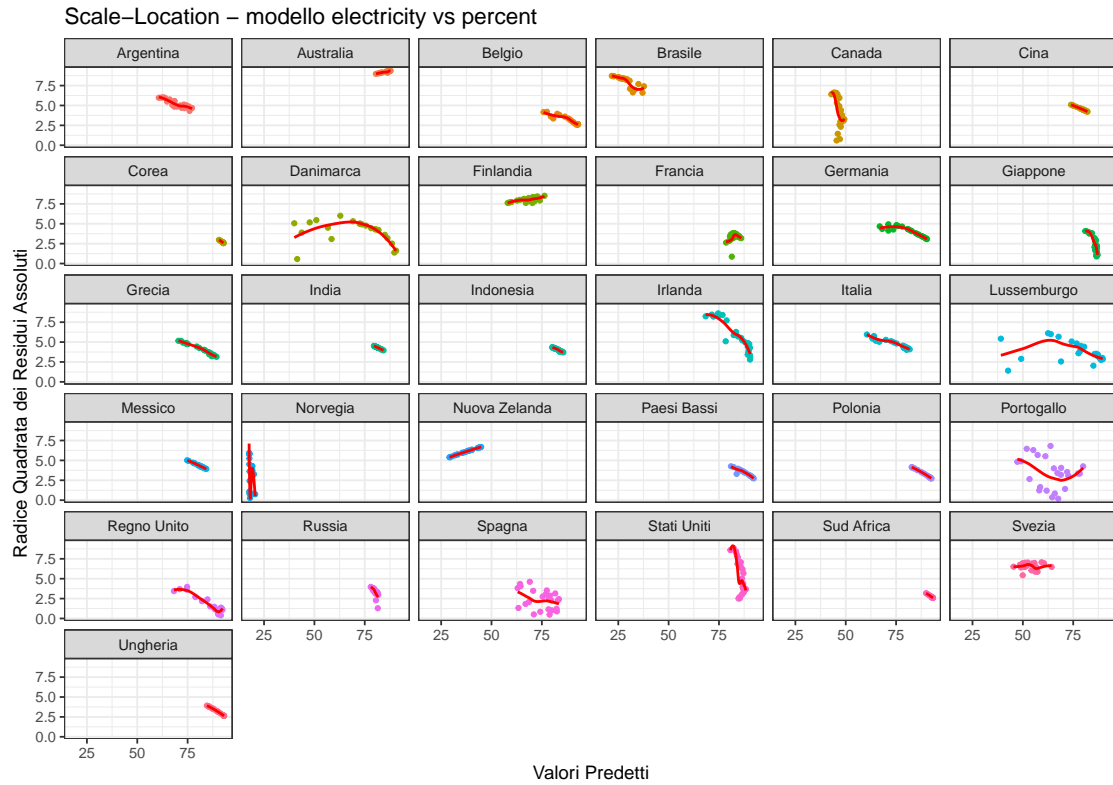


Figura 4.9: Grafico Scale-Location per il modello Elettricità Rinnovabile vs. Percentuale di Popolazione Esposta.

4.4 Energia Rinnovabile e Stress Idrico

In questa sezione, esploriamo la relazione tra la percentuale di energia rinnovabile sul totale dell'energia fornita e lo stress idrico utilizzando un modello di regressione polinomiale di secondo grado. L'obiettivo è indagare se un aumento nella fornitura di energia rinnovabile influisce sullo stress idrico.

4.4.1 Risultati del Modello

Il modello polinomiale è stato applicato con la seguente formula:

$$\text{Stress Idrico} = \alpha + \beta_1 \times (\text{Energia Rinnovabile (\%)}) + \beta_2 \times (\text{Energia Rinnovabile (\%)})^2$$

Dove:

- $\alpha = 25.3974$ è l'intercetta.
- $\beta_1 = -201.8366$ è il coefficiente per il primo termine polinomiale.
- $\beta_2 = 0.6907$ è il coefficiente per il secondo termine polinomiale.

I risultati indicano che il primo termine polinomiale (β_1) è altamente significativo ($p < 2e - 16$), mentre il secondo termine (β_2) non risulta significativo ($p = 0.974$).

Il modello presenta un R^2 di 0.1113, indicando che circa l'11% della varianza nello stress idrico è spiegata dal modello. Questo suggerisce che, sebbene esista una relazione tra energia rinnovabile e stress idrico, una parte significativa della varianza rimane non spiegata.

- **Errore Standard dei Residui:** 20.96
- **F-statistic:** 46.38 su 2 e 741 gradi di libertà, con un p -value $\approx 2.2e - 16$

Gli intervalli di confidenza per i coefficienti del modello sono i seguenti:

- **Intercetta (α):** [23.889, 26.906]
- **Coefficiente per il primo termine polinomiale (β_1):** [-242.977, -160.696]
- **Coefficiente per il secondo termine polinomiale (β_2):** [-40.450, 41.831]

Gli intervalli di confidenza indicano che il primo termine polinomiale ha un impatto negativo sullo stress idrico, mentre il secondo termine non è significativo.

4.4.2 Analisi dei Residui

Grafico dei residui standardizzati Il grafico dei residui standardizzati (Figura 4.10) mostra una distribuzione che evidenzia una certa non linearità, come previsto in un modello polinomiale. Tuttavia, la presenza di outlier potrebbe indicare la necessità di includere ulteriori variabili o di considerare altri modelli per catturare completamente la relazione.

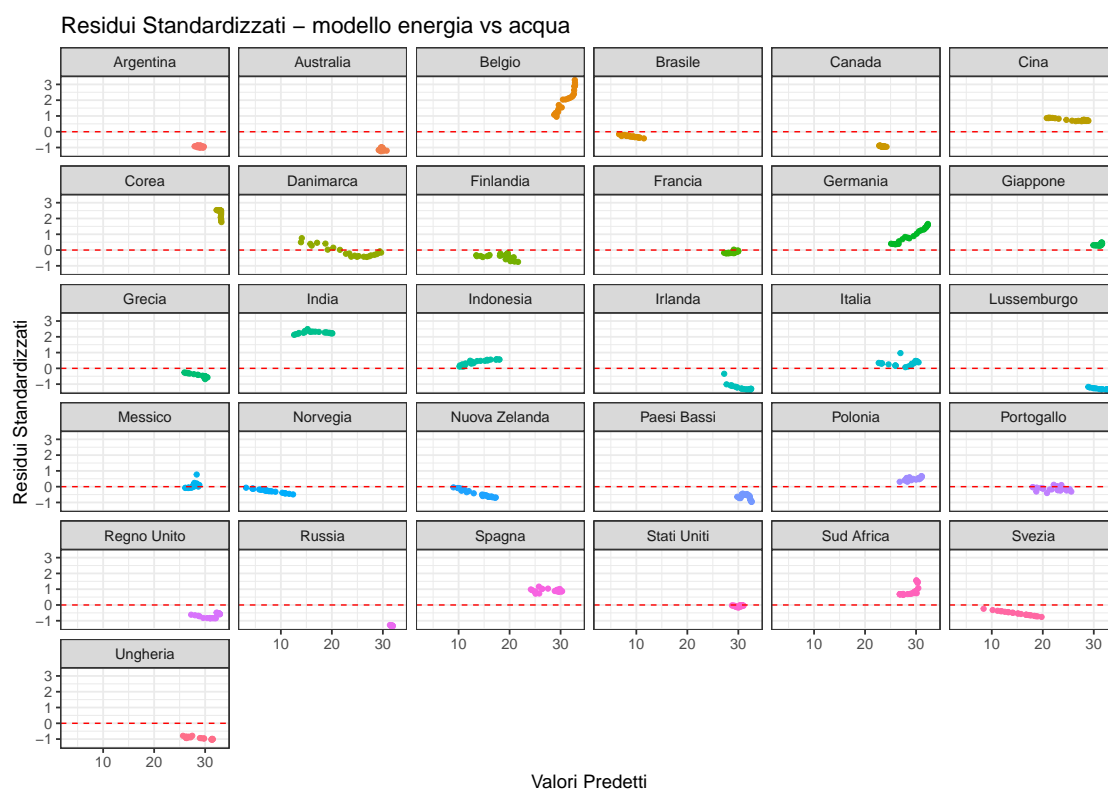


Figura 4.10: Grafico dei Residui Standardizzati per il modello Energia Rinnovabile vs. Stress Idrico.

Grafico Scale-Location Il grafico Scale-Location (Figura 4.11) evidenzia una certa eterogeneità nella varianza dei residui, suggerendo la possibile presenza di problemi di eteroschedasticità nel modello.

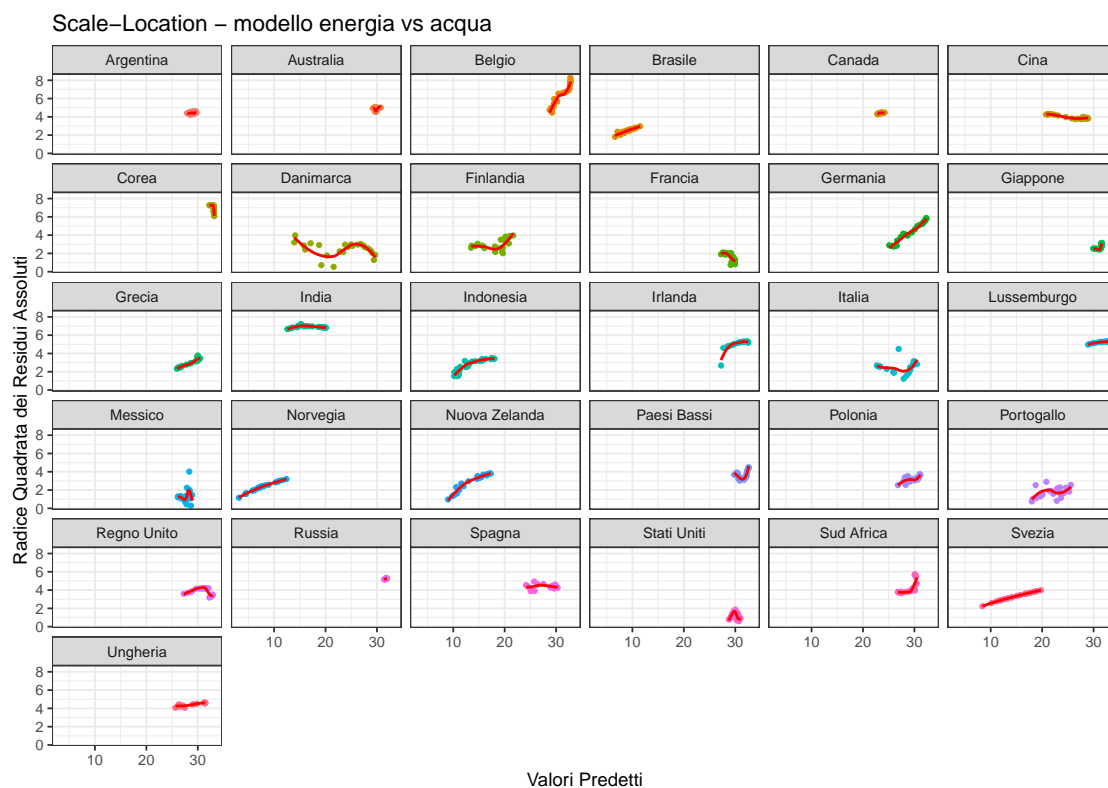


Figura 4.11: Grafico Scale-Location per il modello Energia Rinnovabile vs. Stress Idrico.

QQ Plot Il QQ plot (Figura 4.12) mostra che i residui seguono approssimativamente una distribuzione normale, con alcune deviazioni agli estremi che potrebbero indicare la presenza di outlier.

Il modello suggerisce che esiste una relazione tra l'energia rinnovabile e lo stress idrico, ma è complessa e non completamente catturata da questo modello polinomiale. Il basso valore di R^2 e la non significatività del secondo termine polinomiale indicano che potrebbero esserci altri fattori in gioco. L'analisi dei residui e il QQ plot (Figura 4.12) confermano la presenza di non linearità e possibili outlier, suggerendo la necessità di ulteriori ricerche per migliorare la comprensione di questa relazione.

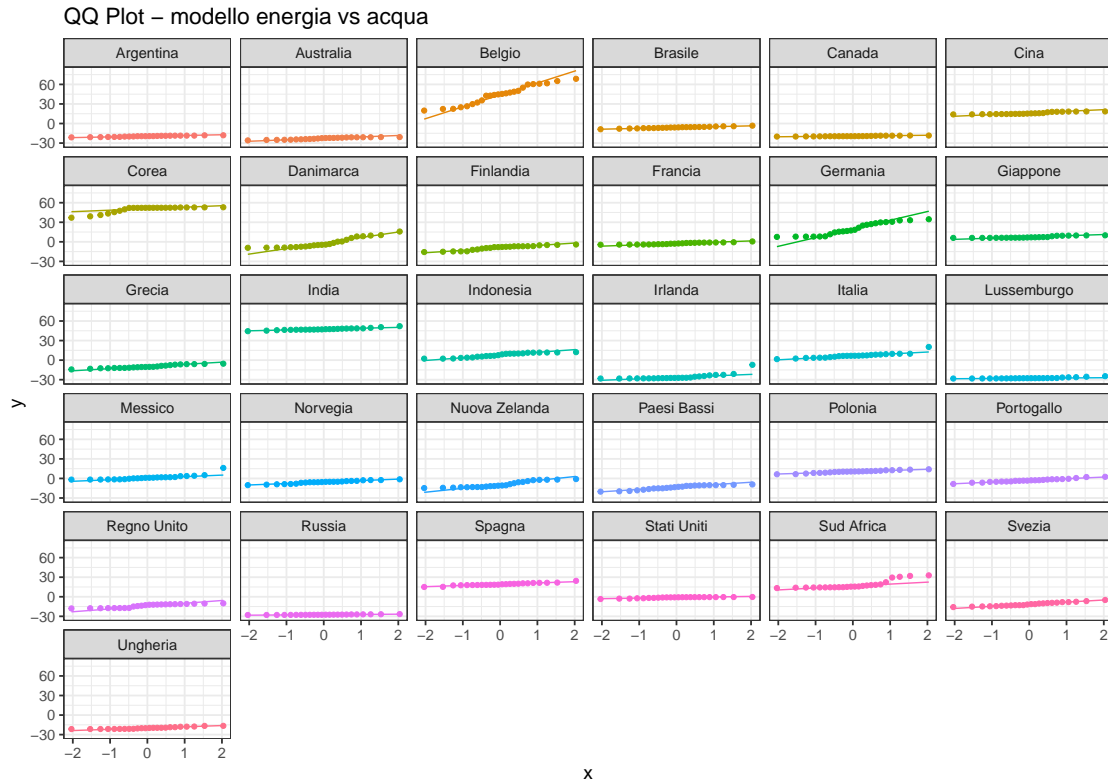


Figura 4.12: QQ Plot per il modello Energia Rinnovabile vs. Stress Idrico.

4.5 Fornitura vs. Elettricità Rinnovabile

Questo modello di regressione lineare esplora la relazione tra la **fornitura di energia rinnovabile** e la **produzione di elettricità da fonti rinnovabili**. L'obiettivo è comprendere come la crescita della fornitura di energia rinnovabile influenzi direttamente la quota di elettricità prodotta da queste fonti.

4.5.1 Risultati e Interpretazione

Il modello di regressione lineare ha prodotto i seguenti risultati:

$$\text{Elettricità Rinnovabile (\%)} = 3.72 + 1.60 \times (\text{Energia Rinnovabile (\%)})$$

- L'intercetta (α) è pari a 3.72 con un errore standard di 0.82. Questo valore rappresenta la percentuale di elettricità rinnovabile quando la fornitura di energia rinnovabile è zero.
- Il coefficiente (β) associato alla fornitura di energia rinnovabile è 1.60 con un errore standard di 0.044, suggerendo che per ogni incremento di un punto percentuale nella fornitura di energia rinnovabile, la produzione di elettricità rinnovabile aumenta di circa 1.60 punti percentuali.

Il coefficiente di determinazione R^2 è pari a 0.638, indicando che circa il 64% della varianza nella produzione di elettricità rinnovabile è spiegata dal modello. Questo suggerisce una forte correlazione tra l'aumento della fornitura di energia rinnovabile e la produzione di elettricità rinnovabile.

Gli intervalli di confidenza per il modello, con un livello di confidenza del 95%, sono i seguenti:

- **Intercetta (α):** [2.11, 5.33]
- **Coefficiente per la Fornitura di Energia Rinnovabile (β):** [1.52, 1.69]

4.5.2 Analisi dei residui

Grafico dei Residui Standardizzati Il grafico dei residui standardizzati (Figura 4.13) mostra che i residui sono distribuiti in modo abbastanza simmetrico attorno allo zero, sebbene vi siano alcuni outlier. Questa distribuzione suggerisce che il modello lineare è un buon fit per i dati, ma indica anche la possibile presenza di altre variabili non incluse nel modello che potrebbero influenzare la produzione di elettricità rinnovabile.

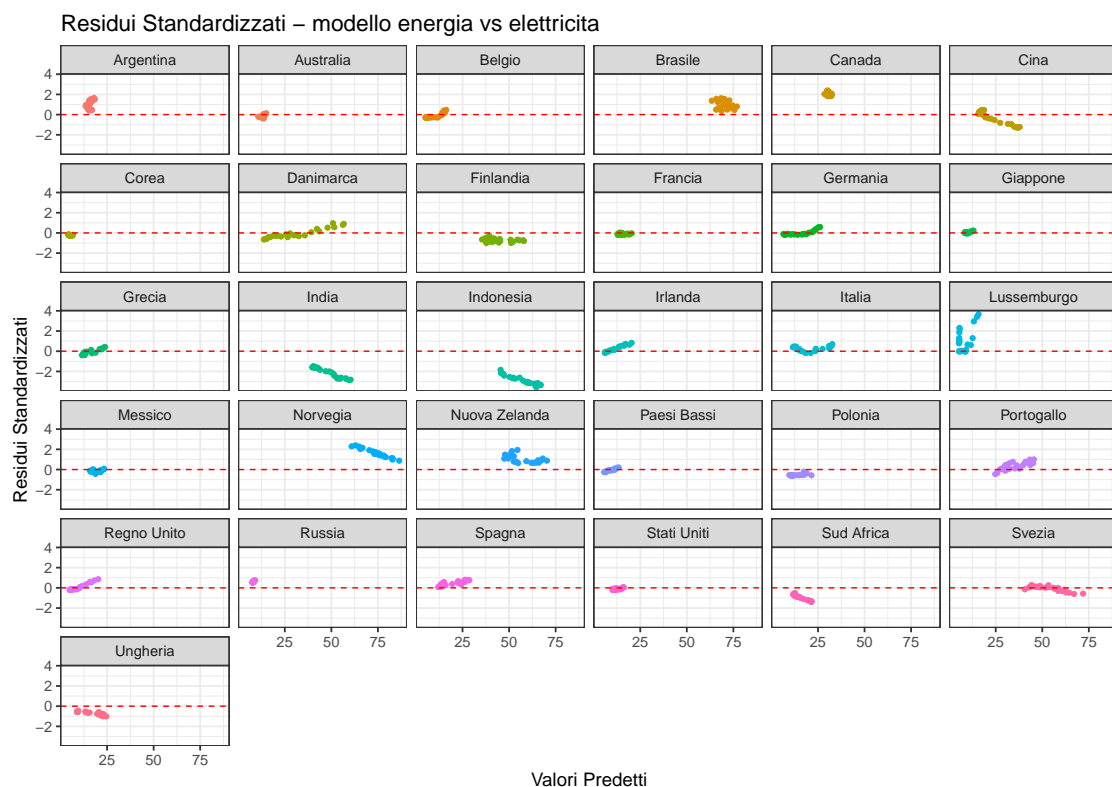


Figura 4.13: Grafico dei Residui Standardizzati per il modello Fornitura di Energia Rinnovabile vs. Elettricità Rinnovabile.

QQ Plot Il QQ plot per il modello (Figura 4.14) suggerisce che i residui si avvicinano abbastanza a una distribuzione normale, con alcune deviazioni agli estremi. Questo indica che, per la maggior parte, il modello di regressione lineare cattura correttamente la distribuzione dei dati, con qualche eccezione.

Grafico Scale-Location Il grafico Scale-Location (Figura 4.15) conferma la distribuzione dei residui standardizzati, mostrando che la varianza dei residui è relativamente costante rispetto ai valori predetti. Tuttavia, alcuni outlier suggeriscono che vi potrebbe essere eteroschedasticità, ovvero variazioni della varianza che potrebbero indicare la presenza di variabili mancanti o non lineari.

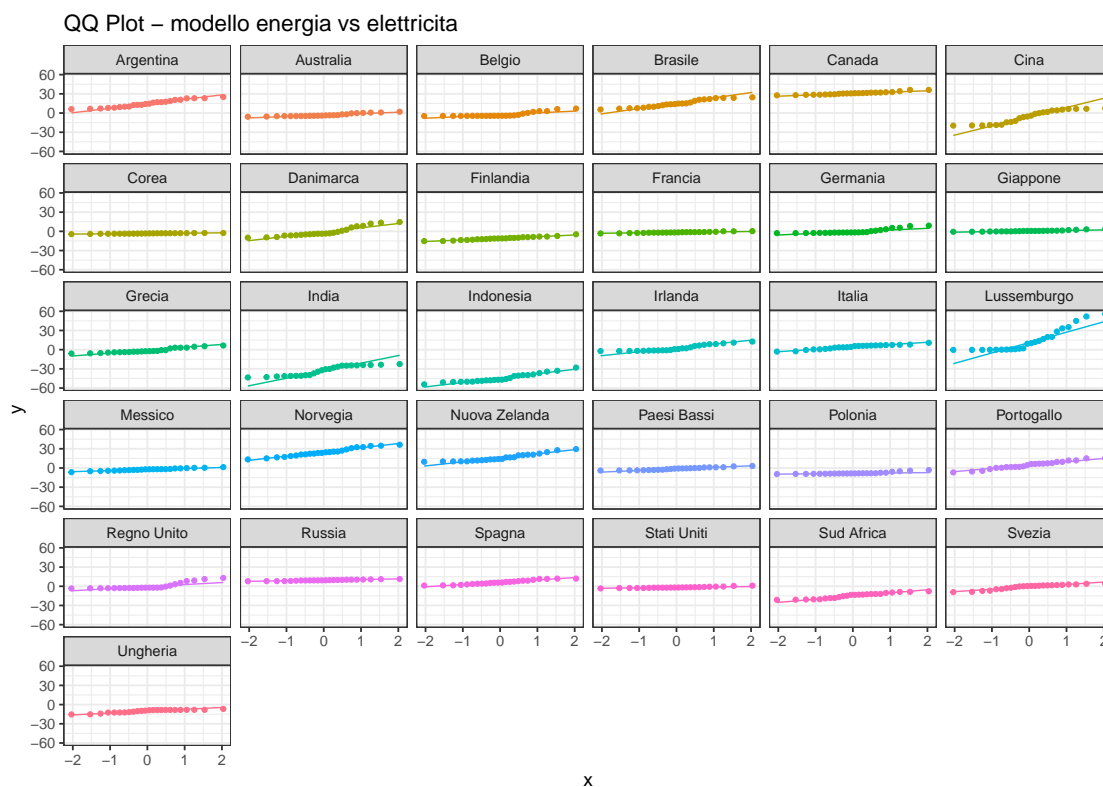


Figura 4.14: QQ Plot per il modello Fornitura di Energia Rinnovabile vs. Elettricità Rinnovabile.

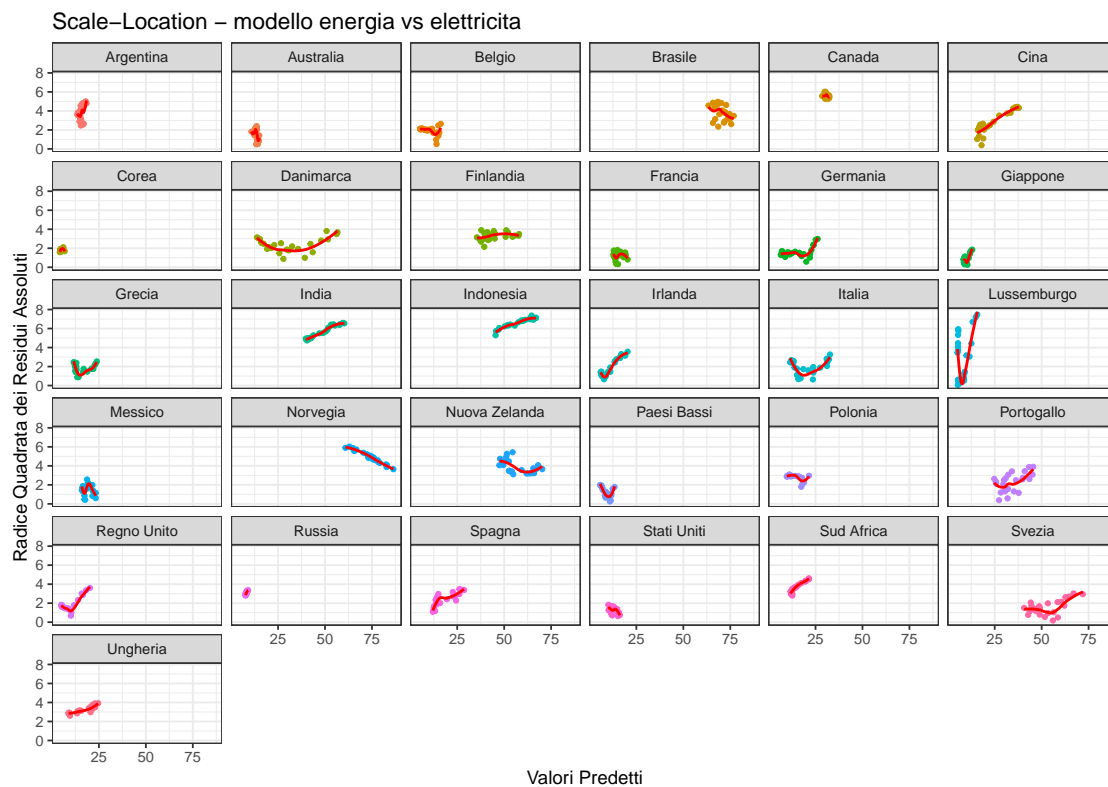


Figura 4.15: Grafico Scale-Location per il modello Fornitura di Energia Rinnovabile vs. Elettricità Rinnovabile.

4.6 Esposizione a PM2.5 e Mortalità

In questa sezione, analizziamo l'impatto combinato dell'esposizione a PM2.5 e dei costi del benessere legati alle mortalità premature sulla mortalità dovuta all'esposizione a PM2.5. L'obiettivo di questa regressione multipla è quantificare come queste due variabili influenzano la mortalità, fornendo così indicazioni utili per le politiche di salute pubblica.

4.6.1 Modello di Regressione Multipla

Il modello di regressione multipla utilizzato è rappresentato dalla seguente equazione:

$$\text{Mortalità da PM2.5} = \alpha + \beta_1 \times \text{Percentuale di Popolazione Esposta} + \beta_2 \times \text{Costi del Benessere}$$

Dove:

- $\alpha = 17.965$ è l'intercetta del modello.
- $\beta_1 = 0.230$ è il coefficiente della percentuale di popolazione esposta a più di 10 microgrammi/m³
- $\beta_2 = 86.117$ è il coefficiente dei costi del benessere espressi in percentuale del PIL.

Il modello suggerisce che per ogni incremento di 1 unità nella percentuale di popolazione esposta, la mortalità dovuta a PIL aumenta di 0.230 unità. Inoltre, un aumento dei costi del benessere di una unità (in percentuale del Prodotto Interno Lordo (PIL)) è associato a un incremento di circa 86 unità nella mortalità da PIL.

4.6.2 Risultati del Modello

Dai risultati della regressione (Tabella 4.1), osserviamo che entrambi i coefficienti β_1 e β_2 sono altamente significativi, con p-value ≤ 0.001 . Il valore di R^2 è 0.9901, indicando che il modello spiega circa il 99% della varianza nella mortalità da PM2.5.

Tabella 4.1: Risultati del Modello di Regressione Multipla

Variabile	Coefficiente	Errore Standard	p-value
Intercetta	17.965	1.918	$\leq 2e-16$
% Popolazione Esposta	0.230	0.029	$7.56e-15$
Costi del Benessere	86.117	0.410	$\leq 2e-16$

Gli intervalli di confidenza al 95% per i coefficienti sono riportati nella Tabella 4.2.

Tabella 4.2: Intervalli di Confidenza al 95% per i Coefficienti del Modello di Regressione Multipla

Variabile	Limite Inferiore	Limite Superiore
Intercetta	14.201	21.729
% Popolazione Esposta	0.173	0.287
Costi del Benessere	85.312	86.922

4.6.3 Analisi dei Residui

Grafico dei Residui Standardizzati Il grafico dei residui standardizzati (Figura 4.16) mostra una distribuzione casuale attorno allo zero, suggerendo che il modello è adeguato e non vi sono evidenti violazioni delle assunzioni di linearità e omoschedasticità.

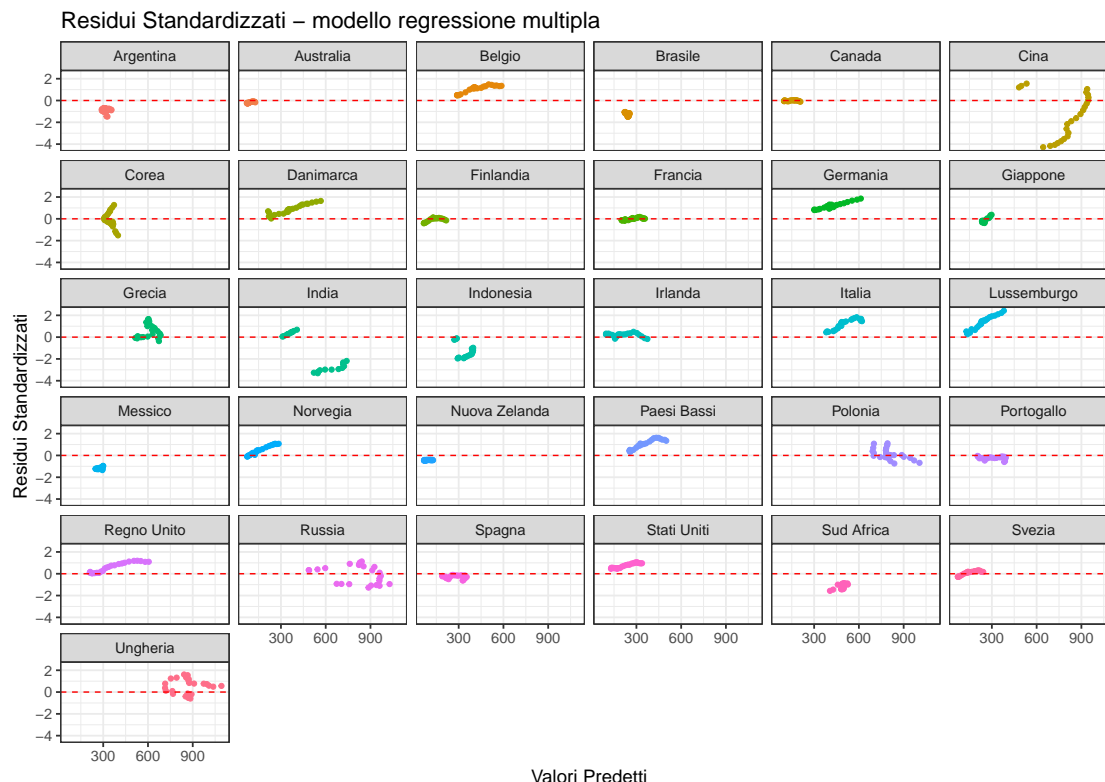


Figura 4.16: Grafico dei residui standardizzati per il modello di regressione multipla.

QQ Plot Il QQ plot (Figura 4.17) mostra che i residui seguono approssimativamente una distribuzione normale, con qualche devianza agli estremi, che potrebbe indicare la presenza di outlier.

Grafico Scale-Location Il grafico Scale-Location (Figura 4.18) conferma che i residui hanno una varianza relativamente costante, supportando l'ipotesi di omoschedasticità.

Questo modello di regressione multipla ha dimostrato una forte capacità predittiva, spiegando quasi la totalità della varianza nella mortalità da PM_{2.5}. I risultati suggeriscono che sia l'esposizione della popolazione a livelli elevati di PM_{2.5} che i costi economici associati alle mortalità premature sono determinanti significativi della mortalità da PM_{2.5}. Queste informazioni sono cruciali per sviluppare politiche di mitigazione che possano ridurre l'impatto sanitario dell'inquinamento atmosferico.

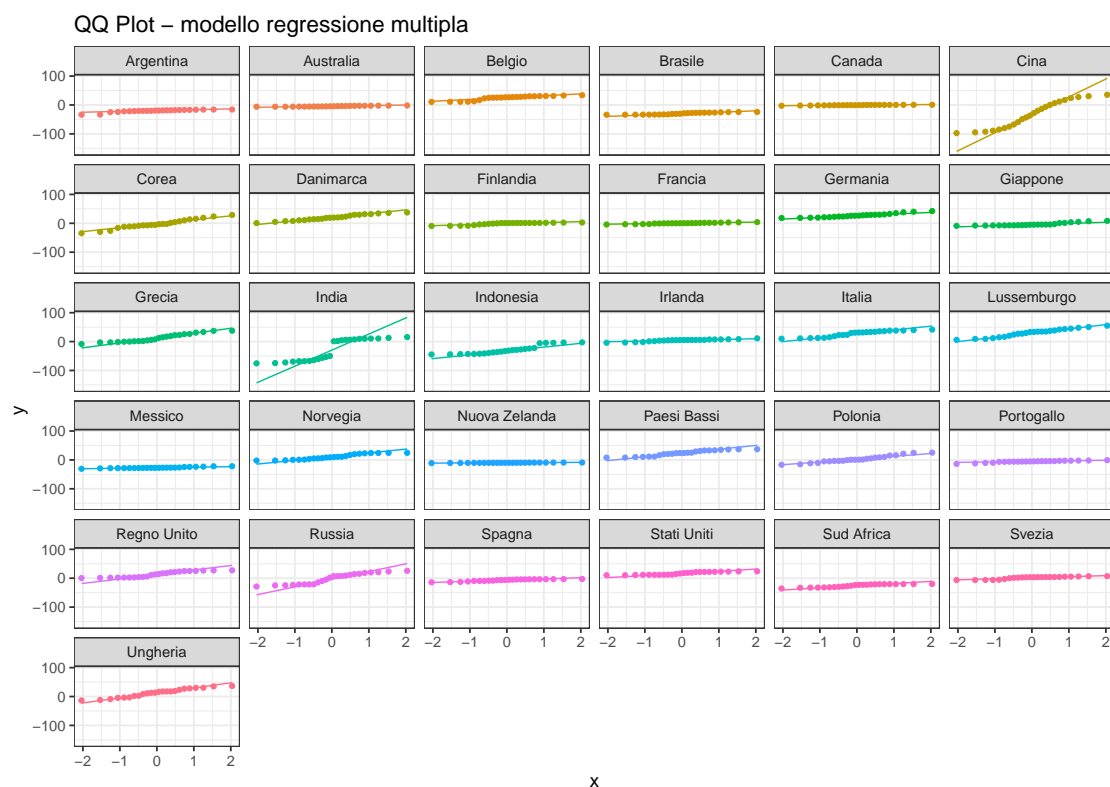


Figura 4.17: QQ plot dei residui per il modello di regressione multipla.

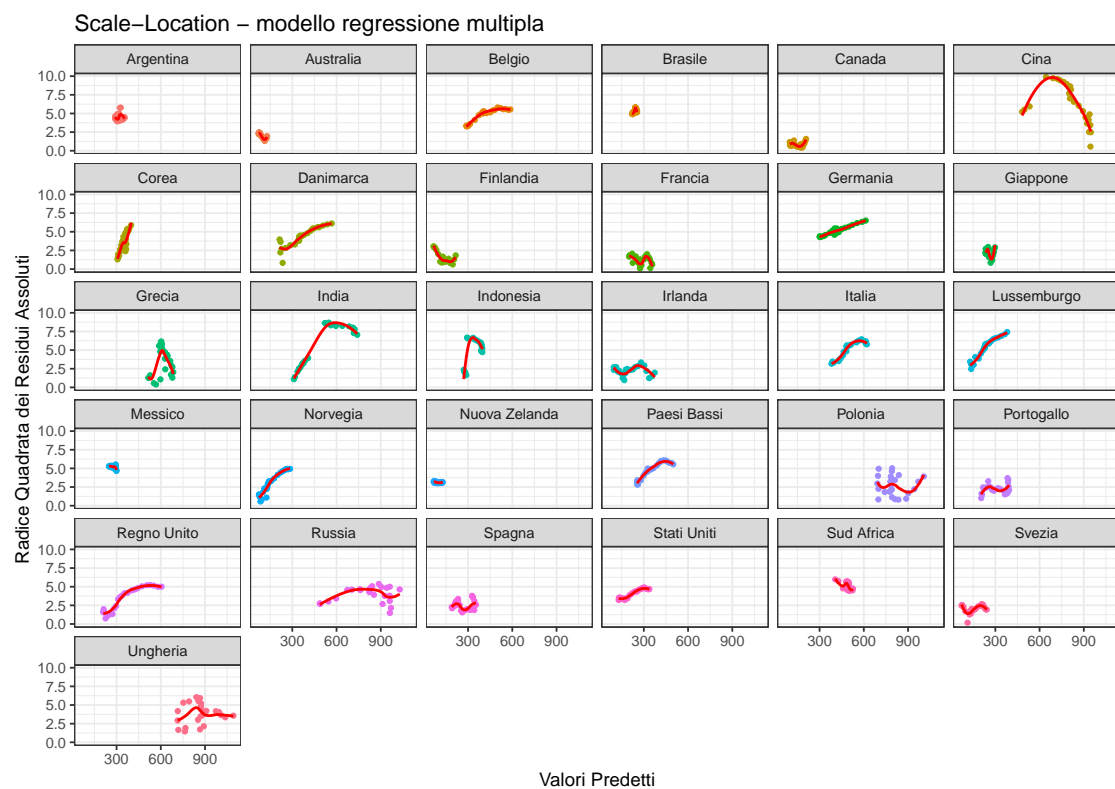


Figura 4.18: Grafico Scale-Location per il modello di regressione multipla.

4.7 Conclusioni e Implicazioni Politiche

I risultati di queste analisi di regressione offrono diverse indicazioni politiche:

- **Riduzione delle Emissioni:** Le forti correlazioni identificate tra la produzione e la domanda di CO₂ indicano che le politiche mirate a ridurre la produzione di CO₂ potrebbero avere un impatto diretto e significativo sulla domanda complessiva di emissioni.
- **Interventi su Energia Rinnovabile e Salute Pubblica:** L'incremento della produzione di energia da fonti rinnovabili sembra ridurre l'esposizione della popolazione a PM2.5, suggerendo che politiche che incentivano l'uso di energia rinnovabile potrebbero anche portare benefici significativi per la salute pubblica.
- **Importanza dei Costi del Benessere:** I costi economici legati alle mortalità premature da PM2.5 sono fortemente correlati con la mortalità stessa, evidenziando la necessità di considerare l'impatto economico e sanitario dell'inquinamento atmosferico nelle decisioni politiche.

Capitolo 5

Analisi dei Cluster

5.1 Introduzione

L'analisi dei cluster è una tecnica di esplorazione dei dati che mira a raggruppare osservazioni simili in cluster omogenei. Questo processo può essere suddiviso in due categorie principali di metodi: metodi non gerarchici e metodi gerarchici.

- **Metodi non gerarchici:** I metodi non gerarchici, noti anche come metodi di partizione, sono utilizzati per suddividere un set di dati in un numero specificato di cluster. Questi metodi sono spesso utilizzati quando si ha un'idea chiara del numero di cluster desiderato o quando si ha una grande quantità di dati che rende impraticabile l'uso dei metodi gerarchici
 1. **K-Means:** Il metodo K-Means è uno dei più popolari. Si basa su un algoritmo che cerca di suddividere i dati in K cluster, dove K è un numero predefinito. L'algoritmo minimizza la somma delle distanze quadrate tra gli elementi di ciascun cluster e il suo centroide.
- **Metodi Gerarchici:** I metodi gerarchici, al contrario, costruiscono una gerarchia di cluster che può essere esplorata a diversi livelli di dettaglio. Questi metodi sono utili quando non si conosce a priori il numero di cluster desiderato o quando si vuole avere una visione completa delle relazioni tra i dati. Ecco alcuni esempi di metodi gerarchici:
 1. **Agglomerativi:** Questi metodi iniziano con ogni osservazione come un cluster separato e poi li uniscono iterativamente in cluster più grandi in base a una metrica di distanza o similarità.
 2. **Divisivi:** Questi metodi iniziano con tutti i dati in un singolo cluster e successivamente li dividono in cluster più piccoli in base alle dissimilarità tra le osservazioni.

La scelta tra metodi non gerarchici e gerarchici dipende dagli obiettivi dell'analisi, dalla natura dei dati e dalla conoscenza a priori del numero di cluster desiderato. Nel prossimo capitolo, esploreremo i dettagli dell'applicazione di ciascun metodo ai nostri dati e analizzeremo i risultati ottenuti.

5.2 Obiettivi dell'Analisi dei Cluster

Nell'ambito di questo studio, l'obiettivo specifico dell'analisi dei cluster è il seguente: **Classificazione dei Paesi in Base alle Politiche Ambientali e Sostenibilità**, identificare raggruppamenti di paesi o regioni che mostrano modelli simili in termini di adozione di politiche ambientali, pratiche di sostenibilità e risultati di crescita verde.

5.3 Selezione delle variabili

La selezione delle variabili è un passaggio cruciale nell'analisi dei cluster. Una scelta accurata delle variabili garantisce che l'analisi sia rilevante, informativa e in grado di produrre intuizioni significative. Inizialmente sono state individuate le seguenti variabili con tematica simile alla classificazione e scartate quelle correlate tra loro.

- Renewable energy supply, % total energy supply
- Environmentally related taxes, % GDP
- Terrestrial protected area, % land area

5.4 Metodi Gerarchici

I metodi gerarchici applicati per l'analisi dei cluster sono stati:

- Metodo del legame singolo
- Metodo del legame completo
- Metodo del legame medio
- Metodo della mediana
- Metodo del centroide

Questi metodi sono stati utilizzati per analizzare la struttura interna dei dati, raggruppando i paesi in cluster omogenei basati su variabili ambientali ed economiche. La matrice delle distanze è stata calcolata utilizzando la **metrica Euclidea**, e i risultati sono stati visualizzati attraverso **dendrogrammi**.

5.4.1 Metodo del legame singolo

Il metodo del legame singolo è uno dei più semplici tra i metodi gerarchici e si basa sulla distanza minima tra due cluster. Applicando questo metodo, sono stati identificati due cluster principali tra i paesi analizzati.

- **Cluster 1:** La maggior parte dei paesi, tra cui **Argentina, Australia, Belgio, Canada, Cina, Corea**, e molti paesi europei come **Francia, Germania e Italia**, appartengono a questo cluster. Questi paesi mostrano caratteristiche simili in termini di variabili analizzate, come una media di tasse ambientali del 2,20% del PIL, una fornitura media di energia rinnovabile del 7,58%, e una percentuale media di area protetta terrestre del 16,50%.

- **Cluster 2:** Questo cluster è composto da un numero minore di paesi, specificatamente **Brasile e Nuova Zelanda**. Questi paesi si distinguono per avere una media inferiore di tasse ambientali (1,84% del PIL), ma una fornitura di energia rinnovabile significativamente più alta (33,86%) e una media di area protetta terrestre del 14,12%.

Il dendrogramma riportato in Figura 5.1 mostra come questi cluster siano stati formati.

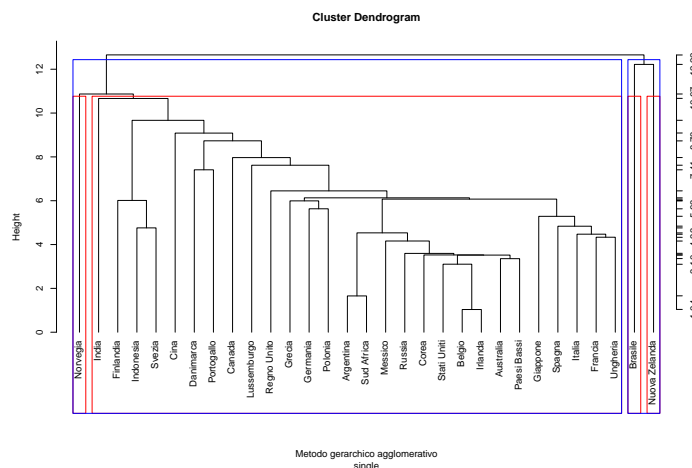


Figura 5.1: Dendrogramma per il metodo del legame singolo

La **non-omogeneità all'interno dei cluster**, indicata da un valore di **74,66**, suggerisce una certa variabilità interna. Tuttavia, la **non-omogeneità tra i cluster** è molto più elevata (**7393,22**), indicando che i cluster sono ben distinti tra loro. Il metodo del legame singolo tende a creare "catene", il che potrebbe spiegare la presenza di un grande cluster e pochi piccoli cluster, come mostrato nello scree plot in Figura 5.2. Il grafico dello scree plot per il metodo evidenzia una curva che scende gradualmente con poche variazioni significative nella pendenza. Tuttavia, si nota un cambiamento intorno a 2-5 cluster, suggerendo che il numero ideale di cluster potrebbe essere in questa gamma, anche se la distinzione tra cluster non è particolarmente marcata rispetto agli altri metodi.

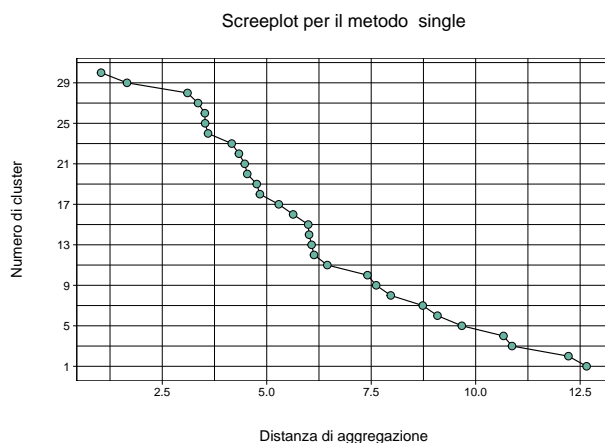


Figura 5.2: Screeplot per il metodo del legame singolo

L'**average silhouette** score di **0,4** suggerisce una moderata coesione interna dei cluster, evidenziando la necessità di esplorare altri metodi per confermare questi risultati.

5.4.2 Metodo del legame completo

Il metodo del legame completo calcola la distanza massima tra due cluster, cercando di minimizzare questa distanza. Questo metodo ha prodotto cluster con una migliore coesione interna rispetto al legame singolo.

- **Cluster 1:** Composto da paesi con una media di tasse ambientali del 2,20% del PIL, una fornitura media di energia rinnovabile del 7,58%, e una percentuale di area protetta terrestre del 16,50%.
- **Cluster 2:** Comprende paesi con una media inferiore di tasse ambientali (1,84% del PIL), ma con una fornitura di energia rinnovabile molto più alta (33,86%) e una percentuale di area protetta terrestre del 14,12%.

Il dendrogramma relativo a questo metodo è mostrato in Figura 5.3.

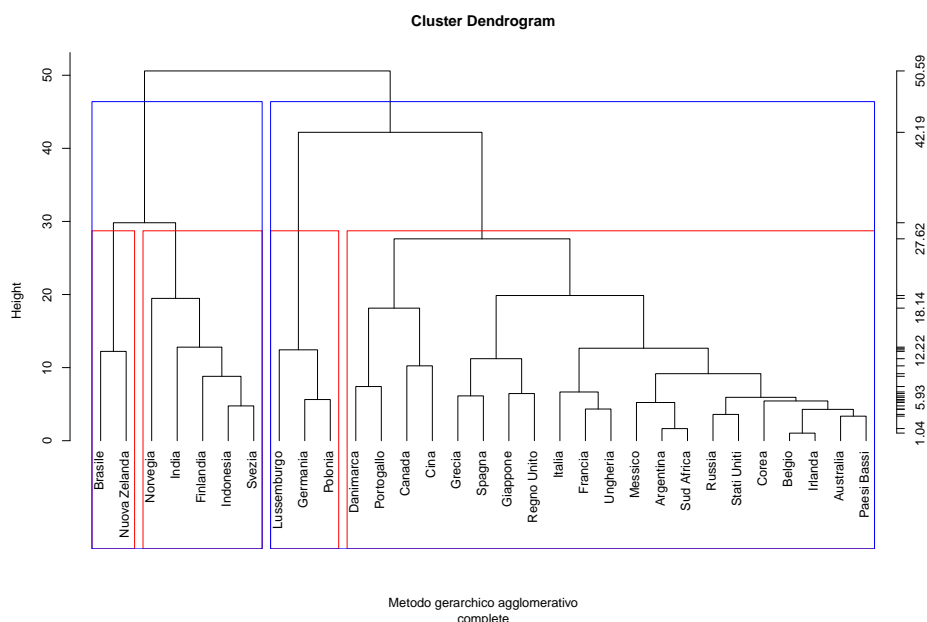


Figura 5.3: Dendrogramma per il metodo del legame completo

La **non-omogeneità all'interno dei cluster** è stata di **144,15**, con una **non-omogeneità tra i cluster** di **7323,73**.

L'**average silhouette score** di **0,51** indica una separazione netta tra i cluster, come evidenziato dallo scree plot in Figura 5.4. Nel grafico dello scree plot relativo al metodo "complete", la curva mostra un cambiamento di pendenza molto marcato dopo 2 cluster, con un graduale appiattimento successivo. Questo indica che il numero ideale di cluster potrebbe essere 2, poiché ulteriori suddivisioni non portano a una riduzione significativa della distanza di aggregazione.

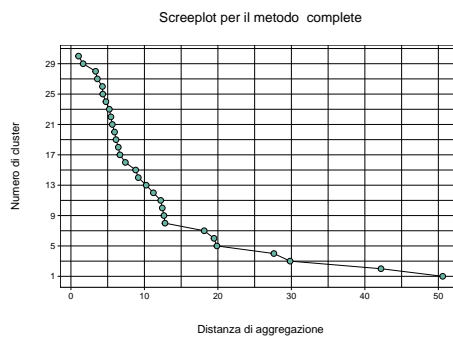


Figura 5.4: Screeplot per il metodo della mediana

5.4.3 Metodo del legame medio

Il metodo del legame medio calcola la media delle distanze tra tutti i punti di due cluster. Questo metodo offre un compromesso tra il legame singolo e il legame completo.

- **Cluster 1:** Paesi con caratteristiche simili a quelle osservate nei metodi precedenti, con una media di tasse ambientali del 2,20% del PIL, una fornitura di energia rinnovabile del 7,58%, e una percentuale di area protetta terrestre del 16,50%.
- **Cluster 2:** Paesi con una media inferiore di tasse ambientali (1,84% del PIL), ma con una fornitura di energia rinnovabile più alta (33,86%) e una percentuale di area protetta terrestre del 14,12%.

Il dendrogramma in Figura 5.5 illustra i cluster formati con questo metodo.

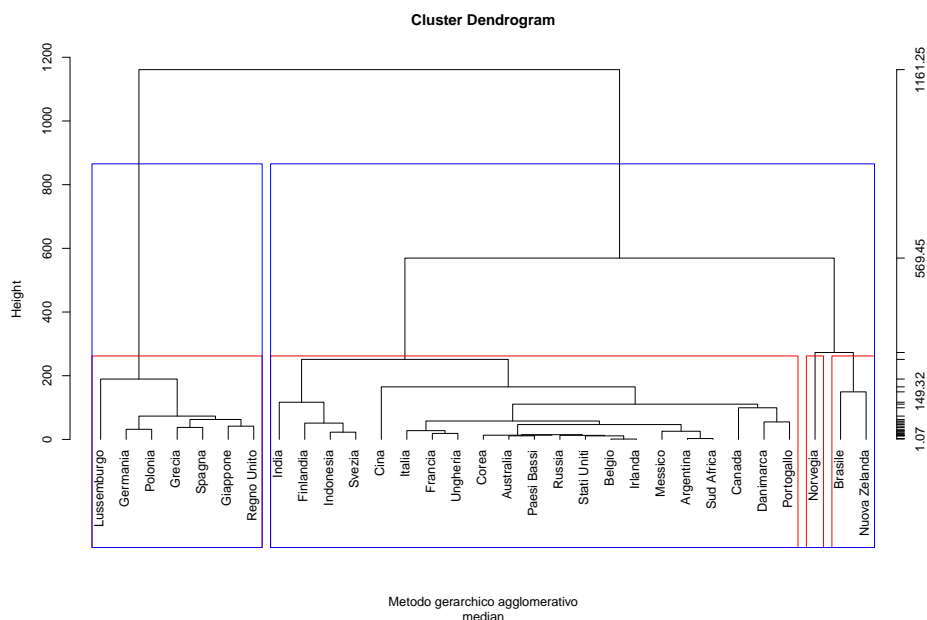


Figura 5.5: Dendrogramma per il metodo del legame medio

La **non-omogeneità interna** di 144,15 e la **non-omogeneità tra i cluster** di 7323,73 sono simili a quelle del legame completo, con un **average silhouette score** di 0,51.

Il grafico dello scree plot in Figura 5.6 per il metodo "median" mostra un calo iniziale piuttosto ripido della distanza di aggregazione, con un cambiamento di pendenza intorno a 3-4 cluster. Questo suggerisce che 3 o 4 cluster potrebbero rappresentare una scelta ideale, in quanto ulteriori suddivisioni portano a riduzioni minime della distanza di aggregazione.

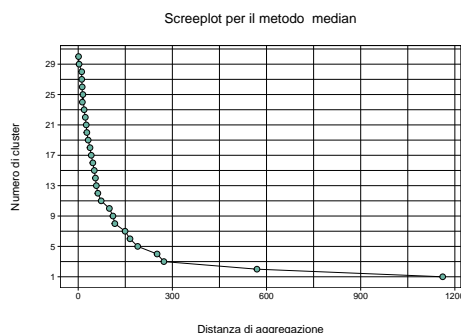


Figura 5.6: Screeplot per il metodo della mediana

5.4.4 Metodo della mediana

Il metodo della mediana, meno comune, calcola la mediana delle distanze tra tutti i punti di due cluster. Questo metodo è particolarmente utile quando la mediana è una metrica significativa.

- **Cluster 1:** Caratterizzato da paesi con una media di tasse ambientali del 2,20% del PIL, una fornitura di energia rinnovabile del 7,58%, e una percentuale di area protetta terrestre del 16,50%.
- **Cluster 2:** Comprende paesi con una media di tasse ambientali inferiore (1,84% del PIL), ma con una fornitura di energia rinnovabile più alta (33,86%) e una percentuale di area protetta terrestre del 14,12%.

Il dendrogramma riportato in Figura 5.7 rappresenta i cluster ottenuti.

Questo metodo ha mostrato una **non-omogeneità interna di 68,51** e una **non-omogeneità tra i cluster di 7399,36**, con un **average silhouette score di 0,51**.

Nel grafico dello scree plot (Figura 5.8), si osserva che la curva mostra un cambiamento significativo nella pendenza intorno ai primi 2-3 cluster. Dopo questo punto, l'aumento della distanza di aggregazione diventa più graduale. Questo suggerisce che un numero ideale di cluster potrebbe essere 3, poiché aggiungere ulteriori cluster oltre questo punto non porta a una significativa riduzione della distanza di aggregazione.

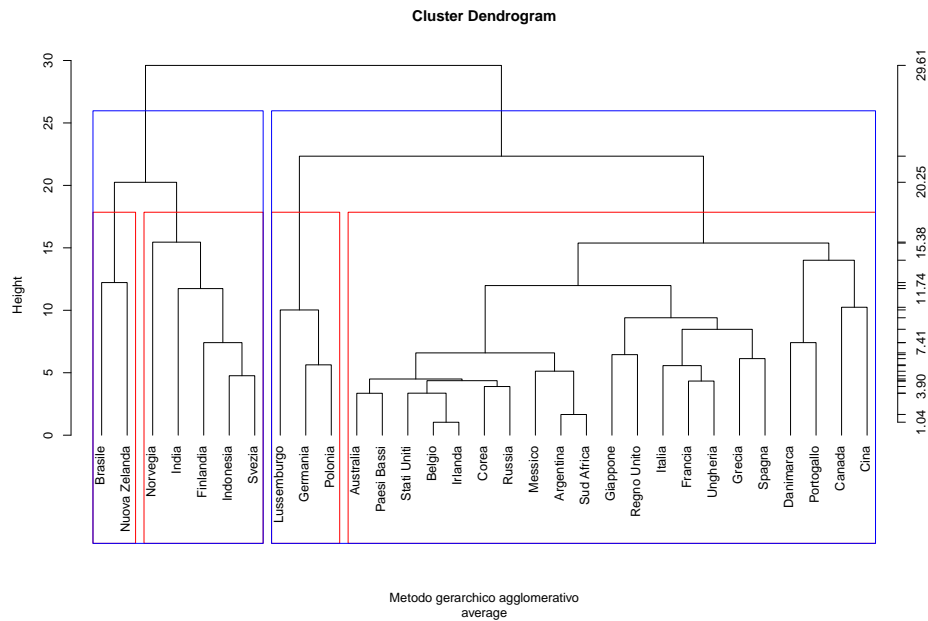


Figura 5.7: Dendrogramma per il metodo della mediana

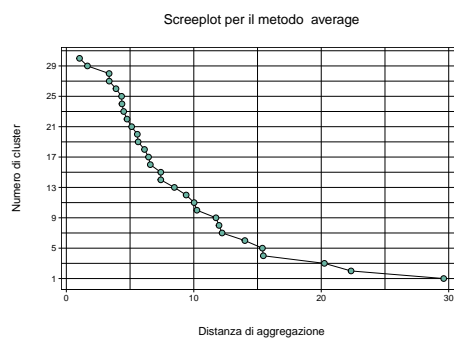


Figura 5.8: Screeplot per il metodo della mediana

5.4.5 Metodo del centroide

Il metodo del centroide utilizza il concetto di centroide, che rappresenta il punto medio di tutti i punti in un cluster. Questo metodo è stato applicato per ottenere una rappresentazione compatta dei cluster.

- **Cluster 1:** Comprende paesi con una media di tasse ambientali del 2,20% del PIL, una fornitura di energia rinnovabile del 7,58%, e una percentuale di area protetta terrestre del 16,50%.
- **Cluster 2:** Paesi con una media inferiore di tasse ambientali (1,84% del PIL), ma con una fornitura di energia rinnovabile più alta (33,86%) e una percentuale di area protetta terrestre del 14,12%.

Il dendrogramma in Figura 5.9 illustra i cluster formati.

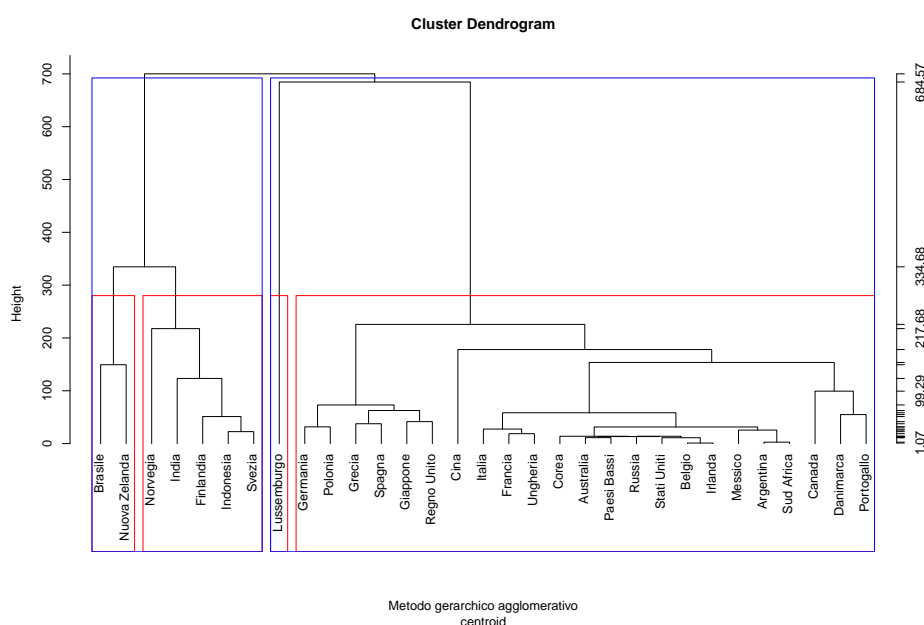


Figura 5.9: Dendrogramma per il metodo del centroide

Il metodo del centroide ha registrato una **non-omogeneità interna di 144,15** e una **non-omogeneità tra i cluster di 7323,73**, ma con un **silhouette score inferiore di 0,27**. Questo indica una separazione meno chiara tra i cluster.

Lo scree plot per il metodo "centroid" (Figura 5.10) presenta una curva che mostra un calo iniziale molto ripido seguito da un appiattimento più graduale. Il punto di inflessione sembra verificarsi intorno a 2-3 cluster, suggerendo che un numero ideale di cluster si colloca in questa gamma. Oltre questo numero, i benefici di suddividere ulteriormente i dati diminuiscono.

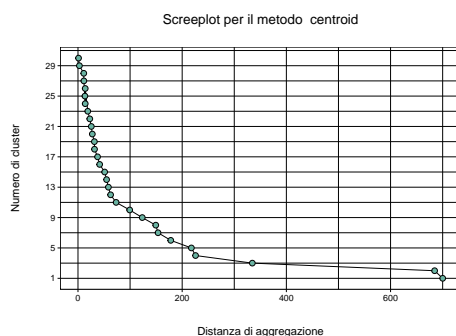


Figura 5.10: Screeplot per il metodo del centroide

5.4.6 Confronto dei Silhouette Score

Il grafico seguente mostra il confronto dei silhouette score per le varie combinazioni di metriche di distanza e metodi di collegamento.

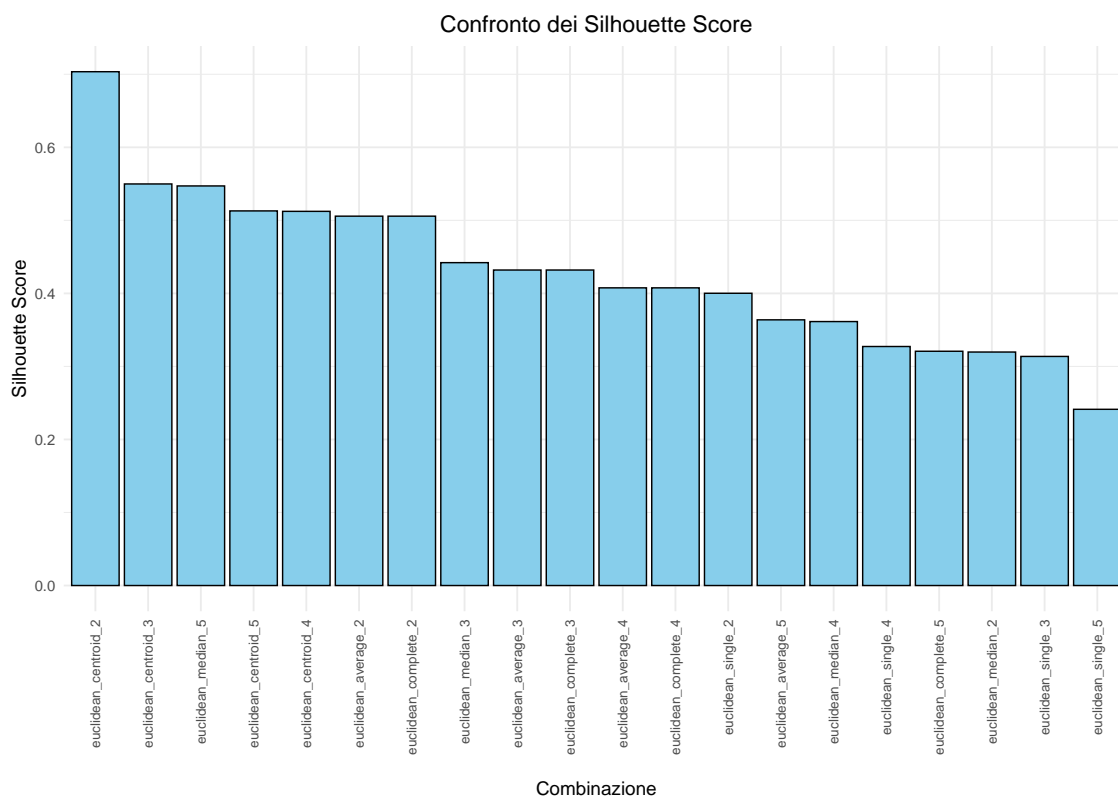


Figura 5.11: Confronto dei Silhouette Score per le diverse combinazioni

Il silhouette score misura la qualità del clustering, riflettendo quanto bene ogni punto dati è assegnato al proprio cluster rispetto ad altri cluster. Valori più alti indicano cluster distinti e ben separati, mentre valori più bassi possono segnalare una sovrapposizione tra cluster.

Dall'analisi del grafico, emerge chiaramente che alcune combinazioni, come il **metodo del legame completo con la distanza euclidea**, hanno prodotto i migliori silhouette score. Questo suggerisce che questi metodi offrono una migliore coesione e distinzione dei cluster. Al contrario,

il **metodo del centroide** ha mostrato performance inferiori, con silhouette score più bassi, indicando una separazione meno netta tra i cluster.

5.4.7 Conclusione sui Metodi Gerarchici di Clustering

L'analisi dei metodi gerarchici di clustering applicata ai dati ambientali ed economici dei paesi ha fornito risultati interessanti e ha messo in luce le caratteristiche distintive dei vari approcci.

Metodo del legame singolo Questo metodo ha evidenziato la presenza di "catene" all'interno dei cluster, con una distinzione meno marcata tra i gruppi rispetto ad altri metodi. Il silhouette score relativamente basso e il graduale calo nello scree plot suggeriscono che questo metodo potrebbe non essere il più adatto per i dati in esame, poiché tende a formare cluster meno distinti.

Metodo del legame completo Questo metodo si è rivelato molto efficace, con un silhouette score elevato e un chiaro gomito nello scree plot a 2 cluster. Questo suggerisce una buona coesione interna e una netta separazione tra i cluster, rendendolo uno dei metodi preferiti per l'analisi dei dati in questione.

Metodo del legame medio e della mediana Entrambi questi metodi hanno mostrato una buona performance, con silhouette score simili e una chiara indicazione di 3-4 cluster come numero ideale. Questi metodi offrono un compromesso tra la coesione interna e la separazione tra i cluster, catturando efficacemente le principali variazioni nei dati.

Metodo del centroide Sebbene il metodo del centroide abbia fornito cluster ragionevolmente compatti, il silhouette score più basso rispetto agli altri metodi indica una separazione meno netta tra i gruppi. Lo scree plot suggerisce che 2-3 cluster rappresentano una scelta ragionevole, ma l'interpretabilità dei cluster potrebbe essere meno chiara.

L'analisi ha dimostrato che la scelta del metodo di clustering e del numero di cluster ideale dipende strettamente dalla natura dei dati e dagli obiettivi specifici dell'analisi. In questo contesto, i metodi del legame completo, del legame medio e della mediana si sono dimostrati i più efficaci, offrendo un buon equilibrio tra coesione interna e distinzione tra i cluster. Il metodo del legame completo, in particolare, ha fornito i risultati più robusti, con un'elevata qualità dei cluster formati.

5.5 Metodi Non Gerarchici

Il metodo K-Means è stato utilizzato per suddividere i dati in un numero specificato di cluster. Questo metodo è stato applicato con diverse configurazioni, esaminando le soluzioni a 2, 3 e 4 cluster per determinare la migliore suddivisione dei paesi in base alle loro caratteristiche ambientali e di sostenibilità.

5.5.1 Risultati K-Means con 2 Cluster

Quando sono stati creati 2 cluster, i risultati hanno indicato una chiara separazione, evidenziando due gruppi distinti di paesi:

- **Cluster 1:** Include **24 paesi** con una media di tasse ambientali del 2,20% del PIL, una fornitura media di energia rinnovabile del 7,58% e una percentuale di area protetta terrestre del 16,50%. Questi paesi mostrano una tendenza moderata verso la sostenibilità ambientale, con valori moderati in tutte le variabili analizzate.
- **Cluster 2:** Comprende **7 paesi** con una media di tasse ambientali inferiore (1,84% del PIL), ma con una fornitura di energia rinnovabile molto più alta (33,86%) e una percentuale di area protetta terrestre del 14,12%. Questi paesi sembrano avere un focus più forte sull'energia rinnovabile rispetto agli altri.

Questa configurazione ha ottenuto un **average silhouette score di 0,51**, suggerendo una separazione netta tra i due cluster. La Figura 5.12 mostra l'allocazione dei paesi in 2 cluster.

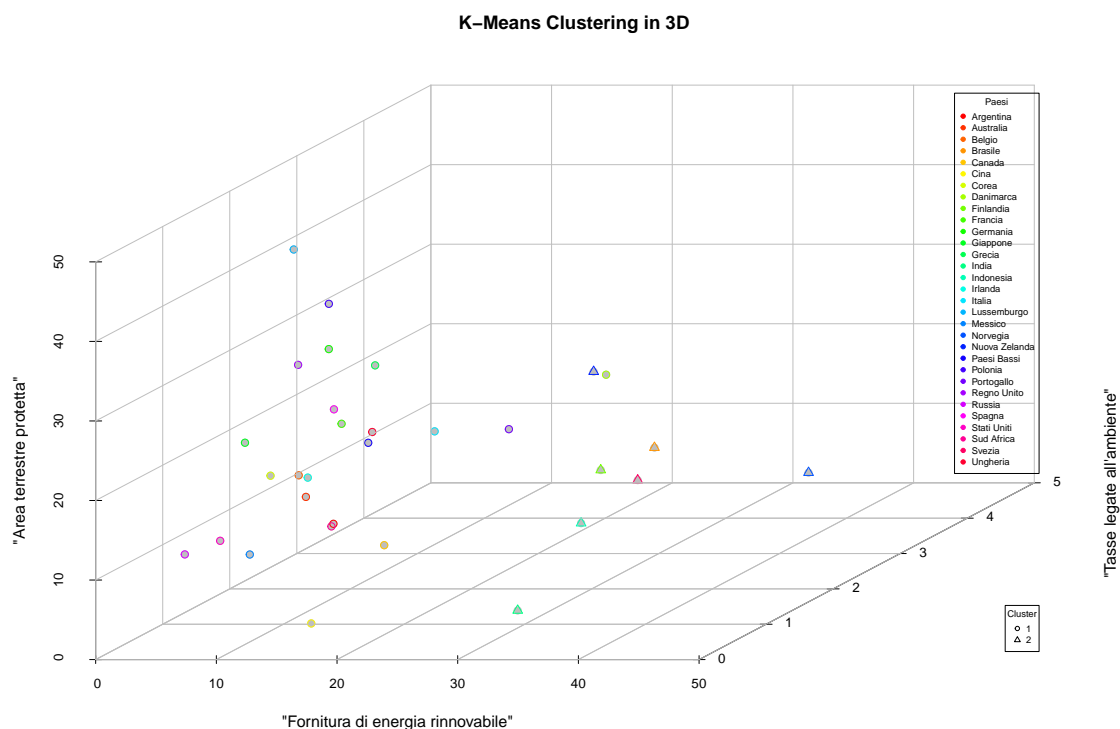


Figura 5.12: Allocazione paesi in 2 cluster

5.5.2 Risultati K-Means con 3 Cluster

- **Cluster 1:** Composto da **6 paesi** con tasse ambientali del 2,36% del PIL, una fornitura di energia rinnovabile del 6,03% e una percentuale di area protetta terrestre del 29,66%. Questi paesi mostrano un impegno significativo nella protezione delle aree terrestri.
- **Cluster 2:** Include **7 paesi** con una fornitura di energia rinnovabile molto elevata (33,86%) e una media di area protetta terrestre di 14,12%, ma con tasse ambientali inferiori (1,84% del PIL). Questo cluster è simile al Cluster 2 nella configurazione a 2 cluster, ma con una maggiore enfasi sull'energia rinnovabile.
- **Cluster 3:** Il cluster più grande, con **18 paesi**, caratterizzati da valori medi di tasse ambientali (2,15% del PIL), fornitura di energia rinnovabile (8,09%) e una percentuale di area protetta terrestre del 12,11%. Questo gruppo rappresenta paesi con politiche ambientali moderatamente equilibrate.

L'average silhouette score per questa configurazione è stato di **0,44**. La Figura 5.13 mostra l'allocazione dei paesi in 3 cluster.

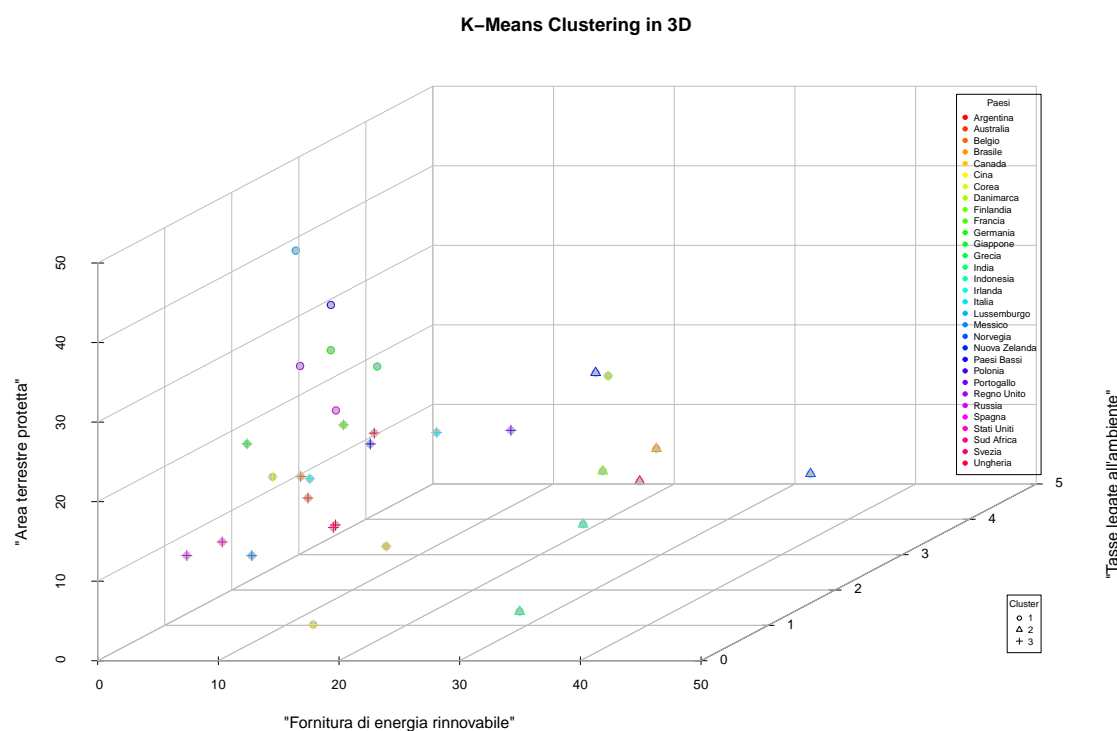


Figura 5.13: Allocazione paesi in 3 cluster

5.5.3 Risultati K-Means con 4 Cluster

La configurazione a 4 cluster ha ulteriormente raffinato la separazione tra i paesi:

- **Cluster 1:** Comprende **6 paesi**, con caratteristiche simili al Cluster 1 della configurazione a 3 cluster, con un'alta percentuale di area protetta.
- **Cluster 2:** Composto da **5 paesi** con la più alta fornitura di energia rinnovabile (36,70%) e una percentuale di area protetta terrestre di 17,49%.
- **Cluster 3:** Include **14 paesi** con valori medi di tasse ambientali (2,10% del PIL) e fornitura di energia rinnovabile (5,67%), con una percentuale di area protetta terrestre del 12,56%. Questi paesi mostrano politiche ambientali bilanciate ma meno incisive rispetto agli altri cluster.
- **Cluster 4:** Contiene **6 paesi** con una fornitura di energia rinnovabile del 19,97% e una percentuale di area protetta terrestre del 8,91%. Questo cluster rappresenta paesi con una moderata adozione di energia rinnovabile e una protezione ambientale più limitata.

Questa configurazione ha ottenuto un **average silhouette score di 0,40**, inferiore alle altre configurazioni, ma ha fornito una separazione più dettagliata, come mostrato in Figura 5.14.

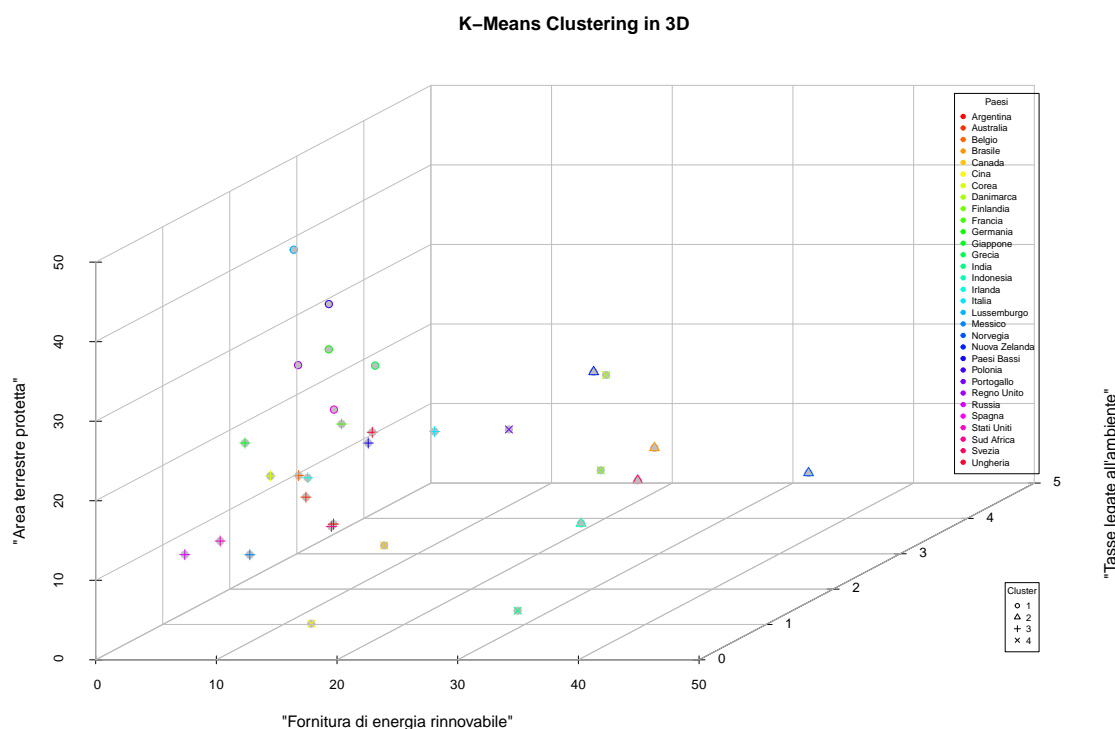


Figura 5.14: Allocazione paesi in 4 cluster

5.5.4 Risultati ottenuti

I risultati ottenuti dall'applicazione del K-Means con configurazioni a 2, 3 e 4 cluster offrono una panoramica interessante su come i paesi si raggruppano in base alle loro caratteristiche ambientali e di sostenibilità. Ogni configurazione ha i suoi punti di forza e debolezza, e la scelta del numero ideale di cluster richiede una riflessione attenta.

Configurazione con 2 Cluster Quando i dati sono stati suddivisi in 2 cluster, è emersa una distinzione molto chiara tra due gruppi di paesi. Il primo cluster raccoglie la maggior parte dei paesi con valori moderati in tutte le variabili analizzate, suggerendo un approccio bilanciato ma non particolarmente aggressivo verso la sostenibilità ambientale. Al contrario, il secondo cluster è caratterizzato da paesi che hanno investito fortemente nell'energia rinnovabile, pur avendo tasse ambientali relativamente basse. Questa configurazione è molto efficace se si desidera ottenere una visione semplice ma chiara delle principali differenze tra i paesi. Il silhouette score di 0,51 conferma che i due cluster sono ben distinti, il che rende questa configurazione particolarmente adatta per chi cerca una separazione netta.

Configurazione con 3 Cluster Passando a 3 cluster, il quadro si arricchisce di sfumature. Qui, i paesi sono stati ulteriormente suddivisi, con un cluster che evidenzia un forte impegno nella protezione delle aree terrestri, un altro che continua a mettere in risalto l'importanza dell'energia rinnovabile, e un terzo che raccoglie la maggior parte dei paesi con politiche ambientali equilibrate ma meno estreme. Anche se l'average silhouette score scende leggermente a 0,44, questa configurazione offre un livello di dettaglio maggiore, che potrebbe essere molto utile per chi desidera esplorare le diverse strategie adottate dai paesi. È una scelta che rappresenta un buon compromesso tra la semplicità della configurazione a 2 cluster e la complessità di quella a 4.

Configurazione con 4 Cluster Infine, la configurazione a 4 cluster ci porta a un livello di dettaglio ancora più alto, suddividendo ulteriormente i paesi in gruppi più specifici. Tuttavia, questo aumento di dettaglio viene ottenuto a scapito della coesione interna dei cluster, come evidenziato dal silhouette score di 0,40. Sebbene questa configurazione permetta di distinguere più chiaramente le diverse strategie ambientali, potrebbe risultare meno utile per chi cerca una visione sintetica e facilmente interpretabile. La maggiore complessità potrebbe rendere difficile trarre conclusioni chiare, e la sovrapposizione tra i cluster potrebbe complicare ulteriormente l'interpretazione dei risultati.

Scelta del Numero Ideale di Cluster La scelta del numero ideale di cluster dipende dagli obiettivi dell'analisi. Se si desidera una visione semplice e chiara delle principali differenze tra i paesi, la configurazione a 2 cluster è probabilmente la scelta migliore. Offre una separazione netta e facilmente interpretabile, ideale per chi cerca un'analisi rapida e immediata. D'altra parte, se l'obiettivo è ottenere una comprensione più dettagliata delle diverse strategie adottate dai paesi in termini di politiche ambientali e sostenibilità, la configurazione a 3 cluster rappresenta un ottimo compromesso. Essa bilancia bene il livello di dettaglio con la coesione interna dei cluster, offrendo una visione più ricca senza perdere di vista l'importanza della distinzione tra i gruppi. La configurazione a 4 cluster, pur offrendo il massimo dettaglio, potrebbe essere troppo complessa per molte applicazioni, a meno che non si abbia una necessità specifica di esplorare le differenze più sottili tra i paesi.

Capitolo 6

Inferenza Statistica

6.1 Stima Puntuale

La stima puntuale consente di stimare il valore di un parametro non noto della popolazione, a partire da un campione osservato. Inizialmente, è necessario determinare il tipo di popolazione in modo da poter applicare il procedimento corretto in relazione alla tipologia di distribuzione. Per ottenere questa informazione, è stato utilizzato il criterio del **chi-quadrato**. In questo contesto, possiamo utilizzare questo criterio per determinare se i dati del campione estratti dall'Italia seguono una distribuzione normale.

6.1.1 Criterio del Chi-Quadrato

Utilizzando il criterio del chi-quadrato (χ^2) sul campione estratto dal dataset, si vuole inizialmente verificare (o rifiutare) l'ipotesi che la distribuzione sia di tipo normale. Le ipotesi di partenza sono così definite:

- **Popolazione di tipo Normale:** H_0
- **Popolazione di tipo Non Normale:** H_1

Il campione preso in esame rappresenta l'**Italia** per un periodo di **24 anni**, e comprende tutte le variabili di interesse che saranno analizzate separatamente. I seguenti indici statistici riassumono i parametri iniziali del campione, inclusi la media campionaria, la deviazione standard e la dimensione del campione, per ciascuna delle variabili considerate:

Variabile	(\bar{x})	(s)	(n)
Production-based CO ₂ emissions	393.80	46.61	24
Demand-based CO ₂ emissions	508.30	61.48	24
Renewable energy supply, % total energy supply	10.48	4.87	24
Renewable electricity, % total electricity generation	24.72	9.24	24
Population with access to improved drinking water sources, % total population	95.63	0.49	24
Water stress, total freshwater abstraction as % total available renewable resources	34.08	5.02	24
Mortality from exposure to ambient PM2.5	532.89	86.83	24
Percentage of population exposed to more than 10 micrograms/m ³	97.33	2.18	24
Welfare costs of premature mortalities from exposure to ambient PM2.5, GDP equivalent	5.40	0.89	24
Environmentally related taxes, % GDP	3.17	0.31	24
Real GDP per capita	38649.05	1633.72	24
Terrestrial protected area, % land area	14.61	3.78	24
Marine protected area, % total exclusive economic zone	4.57	1.45	24

Tabella 6.1: Parametri iniziali (media campionaria, deviazione standard e dimensione del campione) per le variabili analizzate in Italia

Le seguenti tabelle presentano i limiti degli intervalli utilizzati per il test del chi-quadrato, applicato per verificare se i dati campionari delle variabili analizzate seguono una distribuzione normale.

Variabile	Intervallo	Limiti	Frequenze
Production-based CO ₂ emissions	I1	$-\infty, 362.367$	6
	I2	$362.367, 393.802$	4
	I3	$393.802, 425.23$	7
	I4	$425.238, \infty$	7
Valore del Chi-Quadrato Osservato			1
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.023, \chi^2(1 - \alpha/2, df) = 0.001$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.2: Limiti degli intervalli e risultati del chi-quadrato per Italia - Production-based CO₂ emissions

Variabile	Intervallo	Limiti	Frequenze
Demand-based CO ₂ emissions	I1	$-\infty, 466.83$	6
	I2	$466.83, 508.30$	5
	I3	$508.30, 549.77$	7
	I4	$549.77, \infty$	6
Valore del Chi-Quadrato Osservato			0.333
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.024, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.3: Limiti degli intervalli e risultati del chi-quadrato per Italia - Demand-based CO₂ emissions

Variabile	Intervallo	Limiti	Frequenze
Renewable electricity, % total electricity generation	I1	$-\infty, 18.49$	8
	I2	$18.49, 24.72$	7
	I3	$24.72, 30.95$	2
	I4	$30.95, \infty$	7
Valore del Chi-Quadrato Osservato			3.67
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.4: Limiti degli intervalli e risultati del chi-quadrato per Italia - Renewable electricity, % total electricity generation

Variabile	Intervallo	Limiti	Frequenze
Renewable electricity, % total electricity generation	I1	$-\infty, 18.49$	8
	I2	$18.49, 24.72$	7
	I3	$24.72, 30.95$	2
	I4	$30.95, \infty$	7
Valore del Chi-Quadrato Osservato			3.67
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.5: Limiti degli intervalli e risultati del chi-quadrato per Italia - Renewable electricity, % total electricity generation

Variabile	Intervallo	Limiti	Frequenze
Population with access to improved drinking water sources, % total population	I1	$-\infty, 95.29$	9
	I2	$95.29, 95.63$	0
	I3	$95.63, 95.96$	0
	I4	$95.96, \infty$	15
Valore del Chi-Quadrato Osservato			27.00
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 rifiutata: la popolazione non segue una distribuzione normale

Tabella 6.6: Limiti degli intervalli e risultati del chi-quadrato per Italia - Population with access to improved drinking water sources, % total population

Variabile	Intervallo	Limiti	Frequenze
Water stress, total freshwater abstraction as % total available renewable resources	I1	$-\infty, 30.69$	11
	I2	30.69, 34.08	3
	I3	34.08, 37.47	2
	I4	37.47, ∞	8
Valore del Chi-Quadrato Osservato			9.00
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 rifiutata: la popolazione non segue una distribuzione normale

Tabella 6.7: Limiti degli intervalli e risultati del chi-quadrato per Italia - Water stress, total freshwater abstraction as % total available renewable resources

Variabile	Intervallo	Limiti	Frequenze
Mortality from exposure to ambient PM2.5	I1	$-\infty, 474.32$	6
	I2	474.32, 532.89	8
	I3	532.89, 591.45	2
	I4	591.45, ∞	8
Valore del Chi-Quadrato Osservato			4.00
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.8: Limiti degli intervalli e risultati del chi-quadrato per Italia - Mortality from exposure to ambient PM2.5

Variabile	Intervallo	Limiti	Frequenze
Percentage of population exposed to more than 10 micrograms/m ³	I1	$-\infty, 95.86$	3
	I2	95.86, 97.33	5
	I3	97.33, 98.80	12
	I4	98.80, ∞	4
Valore del Chi-Quadrato Osservato			8.33
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 rifiutata: la popolazione non segue una distribuzione normale

Tabella 6.9: Limiti degli intervalli e risultati del chi-quadrato per Italia - Percentage of population exposed to more than 10 micrograms/m³

Variabile	Intervallo	Limiti	Frequenze
Welfare costs of premature mortalities from exposure to ambient PM2.5, GDP equivalent	I1	$-\infty, 4.80$	6
	I2	4.80, 5.40	8
	I3	5.40, 6.01	3
	I4	6.01, ∞	7
Valore del Chi-Quadrato Osservato			2.33
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.10: Limiti degli intervalli e risultati del chi-quadrato per Italia - Welfare costs of premature mortalities from exposure to ambient PM2.5, GDP equivalent

Variabile	Intervallo	Limiti	Frequenze
Environmentally related taxes, % GDP	I1	$-\infty, 2.96$	8
	I2	2.96, 3.17	4
	I3	3.17, 3.38	3
	I4	3.38, ∞	9
Valore del Chi-Quadrato Osservato			4.33
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.11: Limiti degli intervalli e risultati del chi-quadrato per Italia - Environmentally related taxes, % GDP

Variabile	Intervallo	Limiti	Frequenze
Real GDP per capita	I1	$-\infty, 37547.12$	7
	I2	37547.12, 38649.05	5
	I3	38649.05, 39750.98	4
	I4	39750.98, ∞	8
Valore del Chi-Quadrato Osservato			1.67
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.12: Limiti degli intervalli e risultati del chi-quadrato per Italia - Real GDP per capita

Variabile	Intervallo	Limiti	Frequenze
Terrestrial protected area, % land area	I1	$-\infty, 12.06$	8
	I2	$12.06, 14.61$	2
	I3	$14.61, 17.16$	8
	I4	$17.16, \infty$	6
Valore del Chi-Quadrato Osservato			4.00
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.13: Limiti degli intervalli e risultati del chi-quadrato per Italia - Terrestrial protected area, % land area

Variabile	Intervallo	Limiti	Frequenze
Marine protected area, % total exclusive economic zone	I1	$-\infty, 3.59$	4
	I2	$3.59, 4.57$	6
	I3	$4.57, 5.54$	12
	I4	$5.54, \infty$	2
Valore del Chi-Quadrato Osservato			9.33
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 rifiutata: la popolazione non segue una distribuzione normale

Tabella 6.14: Limiti degli intervalli e risultati del chi-quadrato per Italia - Marine protected area, % total exclusive economic zone

Dopo aver applicato il test del chi-quadrato per valutare la normalità delle distribuzioni delle diverse variabili relative all'Italia, sono emersi i seguenti risultati: Variabili che seguono una distribuzione normale:

- *Production-based CO₂ emissions*
- *Demand-based CO₂ emissions*
- *Renewable electricity, % total electricity generation*
- *Mortality from exposure to ambient PM_{2.5}*
- *Welfare costs of premature mortalities from exposure to ambient PM_{2.5}, GDP equivalent*
- *Environmentally related taxes, % GDP*
- *Real GDP per capita*
- *Terrestrial protected area, % land area*

Variabili che **non** seguono una distribuzione normale:

- *Population with access to improved drinking water sources, % total population*
- *Water stress, total freshwater abstraction as % total available renewable resources*
- *Percentage of population exposed to more than 10 micrograms/m³*
- *Marine protected area, % total exclusive economic zone*

6.1.2 Metodo dei Momenti

Il **Metodo dei Momenti** è una tecnica statistica che consente di stimare i parametri di una popolazione, come la media (μ) e la varianza (σ^2), utilizzando i momenti campionari.

Di seguito sono riportati i risultati ottenuti per le variabili analizzate in Italia, applicando il Metodo dei Momenti per stimare μ e σ^2 (Tabella 6.15):

Variabile	μ	σ^2
Production-based CO2 emissions	393.80	2172.15
Demand-based CO2 emissions	508.30	3779.56
Renewable energy supply, % total energy supply	10.48	23.73
Renewable electricity, % total electricity generation	24.72	85.38
Mortality from exposure to ambient PM2.5	532.89	7539.59
Welfare costs of premature mortalities from exposure to ambient PM2.5, GDP equivalent	5.40	0.80
Environmentally related taxes, % GDP	3.17	0.10
Real GDP per capita	38649.05	2669054.00
Terrestrial protected area, % land area	14.61	14.26

Tabella 6.15: Risultati del metodo dei momenti per la stima dei parametri in Italia

Il Metodo dei Momenti fornisce stime puntuali dei parametri μ e σ^2 , che descrivono le caratteristiche centrali e la dispersione delle variabili analizzate. Questi risultati offrono una base quantitativa per comprendere meglio la distribuzione dei dati e per condurre ulteriori analisi statistiche.

Ad esempio, la stima della media per le emissioni di CO_2 basate sulla produzione è **393.80**, mentre la varianza stimata è **2172.15**, indicando una certa variabilità nei dati. Questi parametri ci permettono di caratterizzare il comportamento delle emissioni di CO_2 in Italia durante il periodo analizzato.

6.2 Intervalli di Confidenza

Dopo aver determinato una stima puntuale per i parametri sconosciuti della popolazione, è essenziale estendere questa stima a un intervallo di fiducia. Gli intervalli di confidenza forniscono un range entro il quale il parametro sconosciuto dovrebbe cadere con un certo grado di fiducia, offrendo un'idea della precisione della stima puntuale. In questa sezione, consideriamo un livello di confidenza pari a $1 - \alpha = 0.95$, dove $\alpha = 0.05$. Questo livello rappresenta un buon compromesso tra precisione e affidabilità.

6.2.1 Metodo Pivotal

Il Metodo Pivotal è utilizzato per calcolare gli intervalli di confidenza in popolazioni normali, coprendo diversi scenari. Consideriamo le seguenti casistiche:

- Intervalli di confidenza per μ con σ^2 noto.
- Intervalli di confidenza per μ con σ^2 non noto.
- Intervalli di confidenza per σ^2 con μ noto.
- Intervalli di confidenza per σ^2 con μ non noto.

Intervalli di fiducia di grado $1 - \alpha$ per μ con σ^2 noto

Quando la varianza della popolazione σ^2 è nota, l'intervallo di confidenza per la media μ della popolazione è dato dalla formula:

$$C_n = \bar{X}_n \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},$$

dove:

- \bar{X}_n è la media campionaria,
- $z_{\alpha/2}$ è il valore critico della distribuzione normale standard, pari a 1.96 per $\alpha = 0.05$,
- σ è la deviazione standard della popolazione (conosciuta),
- n è la dimensione del campione.

Per la variabile *Production-based CO₂ emissions* in Italia, con una varianza nota di $\sigma^2 = 2172.15$ e una dimensione campionaria di $n = 24$, l'intervallo di confidenza al 95% per la media è calcolato come segue:

- **Limite inferiore:** $\bar{X}_n - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 393.80 - 1.96 \cdot \frac{\sqrt{2172.15}}{\sqrt{24}} = 375$
- **Limite superiore:** $\bar{X}_n + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 393.80 + 1.96 \cdot \frac{\sqrt{2172.15}}{\sqrt{24}} = 412$

Pertanto, l'intervallo di confidenza al 95% per la media di *Production-based CO₂ emissions* in Italia è (375, 412). La media campionaria di 394 è inclusa in questo intervallo, suggerendo che la stima puntuale è plausibile e ben rappresentativa della media vera della popolazione.

Questo intervallo di confidenza fornisce una misura della precisione della nostra stima della media. Indica che, con un livello di confidenza del 95%, la media vera delle emissioni di CO₂ prodotte in Italia si trova tra 375 e 412. Questo tipo di analisi è essenziale per comprendere meglio l'andamento delle emissioni nel tempo e per confrontare questi risultati con quelli di altri paesi.

Intervalli di fiducia di grado $1 - \alpha$ per μ con σ^2 non nota

Quando la varianza σ^2 della popolazione non è nota, utilizziamo la deviazione standard campionaria S_n per stimare l'intervallo di confidenza. In questo caso, l'intervallo di confidenza per la media μ è calcolato utilizzando la distribuzione t di Student:

$$C_n = \bar{X}_n \pm t_{\alpha/2, n-1} \cdot \frac{S_n}{\sqrt{n}},$$

dove:

- $t_{\alpha/2, n-1}$ è il valore critico della distribuzione t di Student, pari a 2.064 per $\alpha = 0.05$ e $n - 1 = 23$ gradi di libertà,
- S_n è la deviazione standard campionaria.

Per la variabile *Demand-based CO₂ emissions* in Italia, otteniamo:

- **Limite inferiore:** $\bar{X}_n - t_{\alpha/2, n-1} \cdot \frac{S_n}{\sqrt{n}} = 508.30 - 2.064 \cdot \frac{61.48}{\sqrt{24}} = 482$
- **Limite superiore:** $\bar{X}_n + t_{\alpha/2, n-1} \cdot \frac{S_n}{\sqrt{n}} = 508.30 + 2.064 \cdot \frac{61.48}{\sqrt{24}} = 534$

L'intervallo di confidenza al 95% per la media di *Demand-based CO₂ emissions* in Italia è quindi (482, 534), con la media del campione, 508, che rientra in questo intervallo.

In questo caso, l'incertezza legata alla varianza della popolazione si riflette nell'intervallo di confidenza più ampio rispetto al caso precedente. Questo intervallo ci permette di affermare che, con una confidenza del 95%, la media vera delle emissioni di CO₂ basate sulla domanda in Italia si trova tra 482 e 534. Questo è fondamentale per comprendere l'effetto delle politiche economiche e ambientali sul profilo delle emissioni.

Intervalli di fiducia di grado $1 - \alpha$ per σ^2 con μ noto

Quando la media μ della popolazione è nota, l'intervallo di confidenza per la varianza σ^2 può essere calcolato usando la distribuzione chi-quadrato (χ^2). L'intervallo di confidenza è dato da:

$$C = \left(\frac{(n-1) \cdot S_n^2 + n(\bar{X}_n - \mu)^2}{\chi_{\alpha/2, n}^2}, \frac{(n-1) \cdot S_n^2 + n(\bar{X}_n - \mu)^2}{\chi_{1-\alpha/2, n}^2} \right),$$

dove:

- $\chi_{\alpha/2, n}^2$ è il valore critico inferiore della distribuzione χ^2 con n gradi di libertà,
- $\chi_{1-\alpha/2, n}^2$ è il valore critico superiore della distribuzione χ^2 con n gradi di libertà.

Per la variabile *Welfare costs of premature mortalities from exposure to ambient PM_{2.5}, GDP equivalent* in Italia, con una media nota di $\mu = 5.40$, otteniamo:

- **Limite inferiore:** 0.470
- **Limite superiore:** 1.48

Pertanto, l'intervallo di confidenza al 95% per la varianza di *Welfare costs of premature mortalities from exposure to ambient PM_{2.5}, GDP equivalent* in Italia è (0.470, 1.48).

Questo intervallo ci fornisce un range entro cui la varianza vera di questa variabile economica si trova con un grado di confidenza del 95%. Conoscere la varianza è fondamentale per valutare la dispersione dei costi legati alla mortalità prematura e per fare confronti tra paesi o periodi diversi.

Intervalli di fiducia di grado $1 - \alpha$ per σ^2 con μ non noto

Quando né la varianza né la media della popolazione sono note, l'intervallo di confidenza per la varianza σ^2 è calcolato come:

$$C = \left(\frac{(n-1) \cdot S_n^2}{\chi_{\alpha/2, n-1}^2}, \frac{(n-1) \cdot S_n^2}{\chi_{1-\alpha/2, n-1}^2} \right),$$

dove:

- $\chi_{\alpha/2, n-1}^2$ e $\chi_{1-\alpha/2, n-1}^2$ sono i valori critici della distribuzione χ^2 con $n - 1$ gradi di libertà.

Applicando questa formula alla variabile *Real GDP per capita* in Italia, otteniamo:

- **Limite inferiore:** 1.61×10^6
- **Limite superiore:** 5.25×10^6

L'intervallo di confidenza al 95% per la varianza di *Real GDP per capita* in Italia è quindi $(1.61 \times 10^6, 5.25 \times 10^6)$.

Questo intervallo di confidenza fornisce un range di valori plausibili per la varianza del PIL reale pro capite in Italia. Tale intervallo è essenziale per comprendere la variabilità economica tra individui e per pianificare politiche economiche efficaci. Un'alta varianza, ad esempio, potrebbe indicare disuguaglianze economiche significative che richiedono interventi mirati.

6.3 Confronto tra Due Popolazioni

Il confronto tra due popolazioni, in questo caso Italia e Spagna, è fondamentale per comprendere le differenze nelle performance ambientali ed economiche dei due paesi. In particolare, l'analisi si concentra su quattro variabili chiave: le emissioni di CO₂ basate sulla produzione, le emissioni di CO₂ basate sulla domanda, la percentuale di elettricità rinnovabile sul totale generato e il PIL reale pro capite. Di seguito viene descritta la metodologia utilizzata e i risultati ottenuti.

6.3.1 Verifica della Normalità

Prima di procedere al confronto tra le popolazioni italiane e spagnole, è stata condotta una verifica della normalità delle distribuzioni dei dati per ciascuna variabile. La normalità è un presupposto essenziale per l'applicazione corretta del t-test per campioni indipendenti. Per questa verifica è stato utilizzato il test del chi-quadrato, con un livello di significatività $\alpha = 0.05$.

Le ipotesi testate erano le seguenti:

- **Ipotesi Nulla (H_0):** La popolazione segue una distribuzione normale.
- **Ipotesi Alternativa (H_1):** La popolazione non segue una distribuzione normale.

Il campione preso in esame rappresenta la **Spagna** per un periodo di 24 anni. I seguenti indici statistici riassumono i parametri iniziali del campione, inclusi la media campionaria, la deviazione standard e la dimensione del campione, per ciascuna variabile considerata.

Variabile	(\bar{x})	(s)	(n)
Production-based CO ₂ emissions	270.81	35.45	24
Demand-based CO ₂ emissions	318.33	55.49	24
Renewable electricity, % total electricity generation	24.74	9.02	24
Real GDP per capita	33690.05	2921.97	24

Tabella 6.16: Parametri iniziali (media campionaria, deviazione standard e dimensione del campione) per le variabili analizzate in Spagna.

I risultati del test del chi-quadrato sono riportati nelle tabelle seguenti.

Variabile	Intervallo	Limiti	Frequenze
Production-based CO ₂ emissions	I1	$-\infty, 246.904$	7
	I2	$246.904, 270.814$	7
	I3	$270.814, 294.723$	3
	I4	$294.7234, \infty$	7
Valore del Chi-Quadrato Osservato			2.00
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.17: Limiti degli intervalli e risultati del chi-quadrato per Spagna - Production-based CO₂ emissions.

Variabile	Intervallo	Limiti	Frequenze
Demand-based CO ₂ emissions	I1	$-\infty, 280.902$	8
	I2	$280.902, 318.333$	6
	I3	$318.333, 355.763$	5
	I4	$355.763, \infty$	5
Valore del Chi-Quadrato Osservato			1.00
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df) = 5.02, \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.18: Limiti degli intervalli e risultati del chi-quadrato per Spagna - Demand-based CO₂ emissions.

Variabile	Intervallo	Limiti	Frequenze
Renewable electricity, % total electricity generation	I1	$-\infty, 18.651$	7
	I2	$18.651, 24.736$	7
	I3	$24.736, 30.821$	3
	I4	$30.821, \infty$	7
Valore del Chi-Quadrato Osservato			2.00
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df = 5.02), \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.19: Limiti degli intervalli e risultati del chi-quadrato per Spagna - Renewable electricity, % total electricity generation.

Variabile	Intervallo	Limiti	Frequenze
Real GDP per capita	I1	$-\infty, 31719.2$	5
	I2	$31719.2, 33690.0$	5
	I3	$33690.0, 35660.8$	8
	I4	$35660.8, \infty$	6
Valore del Chi-Quadrato Osservato			1
Valori Critici ($\alpha = 0.05$)			$\chi^2(\alpha/2, df = 5.02), \chi^2(1 - \alpha/2, df) = 0.00098$
Conclusione			H_0 accettata: la popolazione segue una distribuzione normale

Tabella 6.20: Limiti degli intervalli e risultati del chi-quadrato per Spagna - Real GDP per capita.

6.3.2 Intervallo di Confidenza per la Differenza tra le Medie

Dopo aver confermato la normalità dei dati, è stato calcolato l'intervallo di confidenza per la differenza tra le medie delle due popolazioni ($\mu_1 - \mu_2$). Questo intervallo fornisce un range entro il quale si prevede che la vera differenza tra le medie possa ricadere con un livello di confidenza del 95%.

L'intervallo di confidenza è stato calcolato utilizzando la formula:

$$x_{n1} - y_{n2} - z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < x_{n1} - y_{n2} + z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Dove:

- x_{n1} e y_{n2} rappresentano le medie campionarie dei campioni italiano e spagnolo.
- σ_1^2 e σ_2^2 sono le varianze dei campioni italiano e spagnolo.
- n_1 e n_2 sono le dimensioni dei campioni.
- $z_{\alpha/2} = 1.96$ è il valore critico della distribuzione normale standard per un livello di confidenza del 95%.

Il calcolo di questi intervalli permette di valutare se le differenze osservate tra le medie delle due popolazioni siano statisticamente significative.

6.3.3 Risultati del Confronto tra le Medie delle Popolazioni

Il calcolo degli intervalli di confidenza permette di valutare se le differenze osservate tra le medie delle due popolazioni, Italia e Spagna, siano statisticamente significative. I risultati sono riportati di seguito per ciascuna delle variabili analizzate.

Emissioni di CO₂ basata sulla produzione Per le emissioni di CO₂ basate sulla produzione, le medie campionarie sono rispettivamente 393.80 per l'Italia e 270.81 per la Spagna. L'intervallo di confidenza al 95% per la differenza tra le medie è dato da:

$$99.56 < \mu_{\text{Italia}} - \mu_{\text{Spagna}} < 146.41$$

Poiché l'intervallo di confidenza non include lo zero, possiamo concludere che la differenza tra le medie è statisticamente significativa, indicando che le emissioni di CO₂ basate sulla produzione sono significativamente maggiori in Italia rispetto alla Spagna.

Emissioni di CO₂ basate sulla domanda Per le emissioni di CO₂ basate sulla domanda, le medie campionarie sono rispettivamente 508.30 per l'Italia e 318.33 per la Spagna. L'intervallo di confidenza al 95% per la differenza tra le medie è dato da:

$$156.83 < \mu_{\text{Italia}} - \mu_{\text{Spagna}} < 223.10$$

Anche in questo caso, l'intervallo di confidenza non include lo zero, suggerendo che le emissioni di CO₂ basate sulla domanda sono significativamente maggiori in Italia rispetto alla Spagna.

Elettricità Rinnovabile, % del Totale Generato Per la percentuale di elettricità rinnovabile generata, le medie campionarie sono molto vicine: 24.72% per l'Italia e 24.73% per la Spagna. L'intervallo di confidenza al 95% per la differenza tra le medie è dato da:

$$-5.19 < \mu_{\text{Italia}} - \mu_{\text{Spagna}} < 5.15$$

In questo caso, l'intervallo di confidenza include lo zero, il che indica che non c'è una differenza significativa tra le medie delle due popolazioni per questa variabile.

PIL Reale Pro Capite Infine, per il PIL reale pro capite, le medie campionarie sono 38649.05 per l'Italia e 33690.05 per la Spagna. L'intervallo di confidenza al 95% per la differenza tra le medie è dato da:

$$3619.67 < \mu_{\text{Italia}} - \mu_{\text{Spagna}} < 6298.32$$

Poiché l'intervallo di confidenza non include lo zero, possiamo concludere che il PIL reale pro capite è significativamente più alto in Italia rispetto alla Spagna.

6.4 Verifica delle Ipotesi

In questa sezione, presentiamo l'analisi di verifica delle ipotesi applicata a due differenti variabili: le emissioni di CO₂ basate sulla produzione e la percentuale di energia rinnovabile nella generazione di elettricità in Italia. Abbiamo eseguito due test di ipotesi, un test unilaterale sinistro per le emissioni di CO₂ e un test unilaterale destro per l'energia rinnovabile. Di seguito discutiamo i dettagli di ciascun test, le ipotesi formulate, e i risultati ottenuti.

6.4.1 Caso 1: Test Unilaterale Sinistro per le Emissioni di CO₂

Tipo di Verifica delle Ipotesi: Questo caso rientra nella **verifica di ipotesi sulla media con varianza nota**. In questo contesto, si assume che la varianza della popolazione sia conosciuta. Di conseguenza, abbiamo utilizzato la statistica Z per eseguire il test di ipotesi, verificando se la media delle emissioni di CO₂ in Italia è inferiore a un valore ipotizzato.

Ipotesi Formulate: Abbiamo testato l'ipotesi che la media delle emissioni di CO₂ basate sulla produzione in Italia sia inferiore a 300 milioni di tonnellate. Le ipotesi sono le seguenti:

- **Ipotesi Nulla (H_0):** La media delle emissioni di CO₂ in Italia è maggiore o uguale a 300 milioni di tonnellate ($\mu \geq 300$).
- **Ipotesi Alternativa (H_1):** La media delle emissioni di CO₂ in Italia è inferiore a 300 milioni di tonnellate ($\mu < 300$).

Risultati: La statistica Z calcolata per questo test è risultata essere 9.859961. Il p-value associato è 1. Pertanto, non rifiutiamo l'ipotesi nulla, concludendo che non ci sono prove sufficienti per affermare che la media delle emissioni di CO₂ in Italia sia inferiore a 300 milioni di tonnellate.

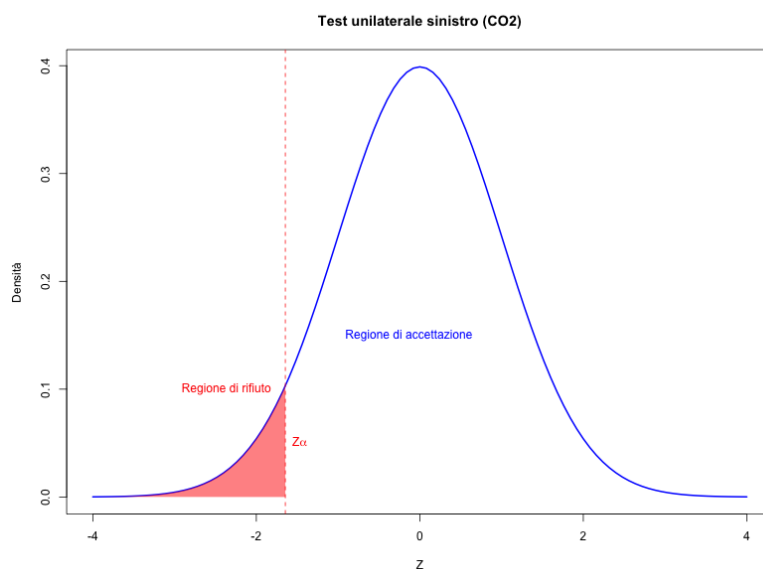


Figura 6.1: Grafico del test unilaterale sinistro per le emissioni di CO₂.

6.4.2 Caso 2: Test Unilaterale Destro per l'Energia Rinnovabile

Tipo di Verifica delle Ipotesi: Anche questo caso rientra nella **verifica di ipotesi sulla media con varianza nota**. In questo test, abbiamo assunto che la varianza della popolazione sia nota e utilizzato la statistica Z per determinare se la media della percentuale di energia rinnovabile nella generazione di elettricità in Italia fosse superiore a un valore ipotizzato.

Ipotesi Formulate: Abbiamo testato l'ipotesi che la media della percentuale di energia rinnovabile nella generazione di elettricità in Italia sia superiore al 50%. Le ipotesi sono le seguenti:

- **Ipotesi Nulla (H_0):** La media della percentuale di energia rinnovabile in Italia è minore o uguale al 50% ($\mu \leq 50\%$).
- **Ipotesi Alternativa (H_1):** La media della percentuale di energia rinnovabile in Italia è superiore al 50% ($\mu > 50\%$).

Risultati: La statistica Z calcolata per questo test è risultata essere -13.40436 . Il p-value associato è 1. Pertanto, non rifiutiamo l'ipotesi nulla, concludendo che non ci sono prove sufficienti per affermare che la media della percentuale di energia rinnovabile in Italia sia superiore al 50%.

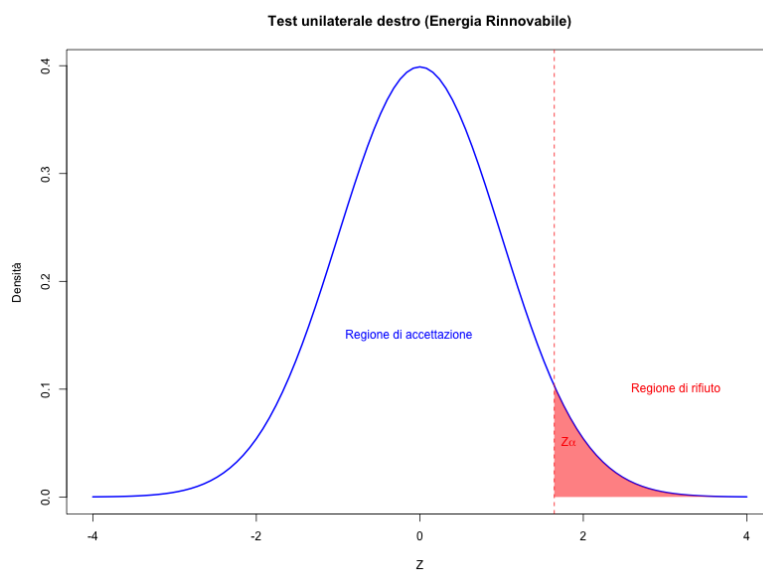


Figura 6.2: Grafico del test unilaterale destro per l'energia rinnovabile.

Capitolo 7

Conclusioni

L'analisi condotta sugli indicatori di crescita verde ha dimostrato come le politiche ambientali rivestano un ruolo fondamentale nel promuovere la sostenibilità economica e ambientale. Attraverso un esame dettagliato dei dati relativi alle emissioni di CO₂, all'uso delle risorse naturali e agli indicatori di salute pubblica, è emerso con chiarezza che le scelte politiche e strategiche dei vari paesi hanno un impatto significativo sia sull'ambiente che sull'economia.

In particolare, l'**analisi temporale delle emissioni di CO₂** (affrontata nel Capitolo 3 - Analisi Esplorativa) ha rivelato tendenze significative nei diversi paesi. Si è evidenziata una correlazione positiva tra l'incremento delle energie rinnovabili e la riduzione delle emissioni, indicando che la transizione verso fonti energetiche sostenibili è non solo possibile, ma anche efficace nel mitigare l'impatto ambientale delle attività produttive.

La **valutazione dell'impatto delle politiche ambientali** (affrontata nel Capitolo 4 - Regressione) ha mostrato che l'adozione di politiche mirate è strettamente associata a variazioni significative nelle emissioni di CO₂ e in altre variabili ambientali. Questi risultati suggeriscono che le politiche ambientali possono essere strumenti potenti per ridurre l'impatto negativo delle attività economiche sull'ambiente.

L'**effetto delle energie rinnovabili sulle emissioni** (affrontato nel Capitolo 4) ha confermato che l'aumento dell'uso di energie rinnovabili porta a una riduzione significativa delle emissioni di CO₂. Questo sottolinea l'importanza di investire in tecnologie verdi e di promuovere politiche energetiche sostenibili.

L'analisi relativa all'**esposizione a PM2.5 e ai costi correlati** (affrontata nel Capitolo 4) ha evidenziato l'impatto negativo dell'inquinamento atmosferico sulla salute pubblica, quantificando i costi economici delle morti premature dovute all'esposizione a PM2.5. Questo rende evidente l'urgenza di politiche che migliorino la qualità dell'aria per ridurre i rischi sanitari e i relativi costi economici.

Il **confronto tra paesi con diverse emissioni di CO₂** (affrontato nel Capitolo 5 - Analisi dei Cluster) ha mostrato che esistono differenze significative nelle emissioni di CO₂ tra paesi con diversi livelli di reddito pro capite. Questi risultati suggeriscono che le politiche ambientali devono essere adattate alle specificità economiche e regionali per essere efficaci.

Da questa analisi possiamo trarre alcune conclusioni fondamentali. Le politiche ambientali ben progettate e implementate sono essenziali per promuovere uno sviluppo economico sostenibile che tenga conto della protezione ambientale. Gli strumenti analitici utilizzati hanno dimostrato che esistono forti interconnessioni tra le variabili economiche e ambientali, e che intervenire su una può avere ripercussioni significative sull'altra.

In particolare, la transizione verso energie rinnovabili e l'adozione di politiche ambientali rigide sono emerse come strategie efficaci per ridurre le emissioni di CO₂ e migliorare la salute pubblica. Tuttavia, è cruciale bilanciare queste politiche con le esigenze economiche per evitare effetti negativi sullo sviluppo e garantire che i benefici ambientali non vadano a scapito della crescita economica.

Gli insight ottenuti da questa analisi possono guidare i decisori politici e le imprese nella formulazione di strategie che promuovano la crescita economica, proteggano le risorse naturali e migliorino il benessere delle generazioni future. È essenziale continuare a monitorare e adattare queste politiche in risposta alle sfide emergenti, assicurando così un futuro sostenibile e prospero per tutti.

Glossario

- Biodiversità** La varietà di vita sulla Terra in tutte le sue forme, compresa la diversità genetica, delle specie e degli ecosistemi
- Boxplot** Un grafico a scatola che rappresenta la distribuzione statistica dei dati
- Correlazione** La correlazione è una misura statistica che descrive la relazione tra due variabili. Un'alta correlazione positiva significa che quando una variabile aumenta, anche l'altra tende ad aumentare, mentre una correlazione negativa indica che quando una variabile aumenta, l'altra tende a diminuire
- Distribuzione dei Dati** La distribuzione dei dati descrive come i dati sono sparsi o distribuiti su un range di valori. Può includere informazioni su media, mediana, varianza e altri aspetti statistici
- Emissioni territoriali** Le emissioni territoriali si riferiscono alle emissioni di gas serra che avvengono entro i confini di un paese
- Energiewende** Un ambizioso piano di transizione energetica adottato dalla Germania per ridurre le emissioni di carbonio e promuovere l'uso di energie rinnovabili
- Imputazione Mediana** Metodo di sostituzione di dati mancanti con la mediana dei valori disponibili
- Interpolazione Lineare** Metodo di stima di valori mancanti in una serie di dati, assumendo che i cambiamenti tra i punti noti siano lineari
- Outlier** Un outlier è un punto dati che si discosta significativamente dagli altri punti del dataset. Gli outlier possono indicare variabilità nel sistema, errori sperimentali, o dati eccezionali che meritano un'attenzione particolare
- Politiche Ambientali** Insieme di regolamenti e azioni adottate da governi o organizzazioni per gestire le risorse naturali e ridurre l'impatto ambientale delle attività umane
- sostenibilità ambientale** Capacità di un sistema di soddisfare le proprie esigenze attuali senza compromettere la capacità delle generazioni future di soddisfare le proprie
- Stress Idrico** Situazione in cui la domanda di acqua supera la disponibilità durante un determinato periodo o quando la sua qualità limita l'uso dell'acqua
- Valore di una Vita Statistica** Il Valore di una Vita Statistica (VSL) è una misura economica utilizzata per valutare il valore monetario della vita umana, spesso utilizzata in analisi costi-benefici per politiche e interventi di sicurezza e salute pubblica

Acronimi

AIE	Agenzia Internazionale dell'Energia
CO ₂	Diossido di Carbonio
EDA	Exploratory Data Analysis
FWA	Freshwater Abstractions
IEA	International Energy Agency
IO	Input-Output
IQR	Intervallo Interquartile
IUCN	Unione Internazionale per la Conservazione della Natura
JMP WSS	Joint Monitoring Programme for Water Supply and Sanitation
ML	Machine Learning
OCSE	Organizzazione per la Cooperazione e lo Sviluppo Economico
OECD	Organization for Economic Co-operation and Development
OMS	Organizzazione Mondiale della Sanità
PIL	Prodotto Interno Lordo
PM2.5	Particulate Matter 2.5
Q1	Primo Quartile
Q3	Terzo Quartile
SI	Stress Idrico
TES	Totale Fornitura Energetica
UNDESA	Organizzazione delle Nazioni Unite per lo Sviluppo Sostenibile
UNICEF	Fondo delle Nazioni Unite per l'Infanzia
USD	United States Dollar
VSL	Valore di una Vita Statistica
WHO	World Health Organization
ZEE	Zona Economica Esclusiva

Bibliografia

Rapporto dell'UNEP sulla valutazione economica dei benefici della protezione ambientale (2020)
Rapporto del World Economic Forum sul Global Competitiveness Report (2022) Rapporto del McKinsey Global Institute sulla trasformazione sostenibile https://data.worldbank.org/indicator/ER.LND.PTLD.ZS?locations=TG&name_desc=false