



UNIVERSITÀ DEGLI STUDI DI SALERNO

Dipartimento di Informatica

Corso di Statistica e Analisi dei Dati

# **Analisi sul Dataset:** *Labour Force Statistics*

AUTORE

**Giovanni Borrelli**

Matricola: 0522501807

Anno Accademico 2023-2024

---

## Indice

---

<b>1</b>	<b>Introduzione</b>	<b>1</b>
1.1	Definizione di Forza Lavoro . . . . .	1
1.2	Descrizione del dataset . . . . .	1
1.3	Obiettivo dell'analisi . . . . .	2
1.4	Accesso al dataset . . . . .	2
<b>2</b>	<b>Serie temporali e Scatterplot</b>	<b>3</b>
2.1	Serie Temporalì . . . . .	3
2.2	Scatterplot e Regressione Lineare . . . . .	29
<b>3</b>	<b>Barplot</b>	<b>52</b>
3.1	Barplot sul sesso . . . . .	52
3.2	Barplot sull'età . . . . .	93
<b>4</b>	<b>Istogrammi e Boxplot</b>	<b>110</b>
4.1	Istogrammi . . . . .	110
4.2	Boxplot . . . . .	121
<b>5</b>	<b>Cluster Gerarchico</b>	<b>122</b>
5.1	Cluster Gerarchico . . . . .	122
5.2	Analisi sui cluster . . . . .	131

---

<b>6</b>	<b>Cluster K-MEANS</b>	<b>135</b>
6.1	Cluster K-MEANS . . . . .	135
6.2	Analisi sui cluster . . . . .	140
<b>7</b>	<b>Variabili Aleatorie</b>	<b>143</b>
7.1	Scelta della funzione di distribuzione . . . . .	143
7.2	Stima puntuale e Intervalli di Confidenza . . . . .	146
7.3	Verifica delle ipotesi . . . . .	148
7.4	Test del chi quadrato . . . . .	149
	<b>Bibliografia</b>	<b>150</b>

# CAPITOLO 1

---

## Introduzione

---

### 1.1 Definizione di Forza Lavoro

La Forza Lavoro si riferisce al numero totale di persone che sono disponibili per svolgere lavoro in un'area geografica specifica, in un settore industriale o in un'azienda. Questo termine è spesso utilizzato in economia e nel contesto dell'occupazione per descrivere la popolazione di lavoratori attivi. La forza lavoro comprende sia le persone occupate che quelle disoccupate che sono in cerca di impiego. L'analisi della forza lavoro è fondamentale per comprendere l'economia di una regione o di un paese, nonché per valutare la disponibilità di risorse umane in diversi settori. I tassi di occupazione, disoccupazione e partecipazione alla forza lavoro sono indicatori comuni utilizzati per valutare la dinamica della forza lavoro.

### 1.2 Descrizione del dataset

La tabella "LFS by sex and age" contiene le statistiche per 38 paesi, dal 2010 al 2022: sull'occupazione, disoccupazione, forza lavoro e popolazione. Tali dati sono suddivisi sia per sesso (maschio o femmina) che per categoria di età (15-19 anni,

20-24 anni, ecc...). Ecco un elenco di tutti i paesi: Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, OECD Countries, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkiye, United Kingdom, United States. I valori del paese "OECD Countries" corrispondono alla somma dei valori dei 38 paesi.

## 1.3 Obiettivo dell'analisi

L'obiettivo di questa analisi sul dataset è visualizzare e comprendere le differenze sulla forza lavoro di ogni paese, cercando di capirne le motivazioni. Analizzando sia le differenze della forza lavoro sul sesso che sull'età, si può avere anche una panoramica più completa di come il mondo del lavoro si sia evoluto dal 2010 fino al 2022. Si possono fare anche confronti importanti avendo a disposizione i dati della disoccupazione e della popolazione di un paese, che possono aiutare a scoprire le problematiche che si celano dietro ogni nazione.

## 1.4 Accesso al dataset

Per poter utilizzare il dataset su R, basta andare sul sito <https://stats.oecd.org/>, nella sezione "Labour", poi cliccare su "Labour Force Statistics", e selezionare la tabella "LFS by SEX and AGE". A questo punto, è possibile selezionare il pulsante "Export" e scegliere il formato output desiderato. In questo caso, è stato scelto un file .csv che è stato successivamente caricato su R Studio.

## CAPITOLO 2

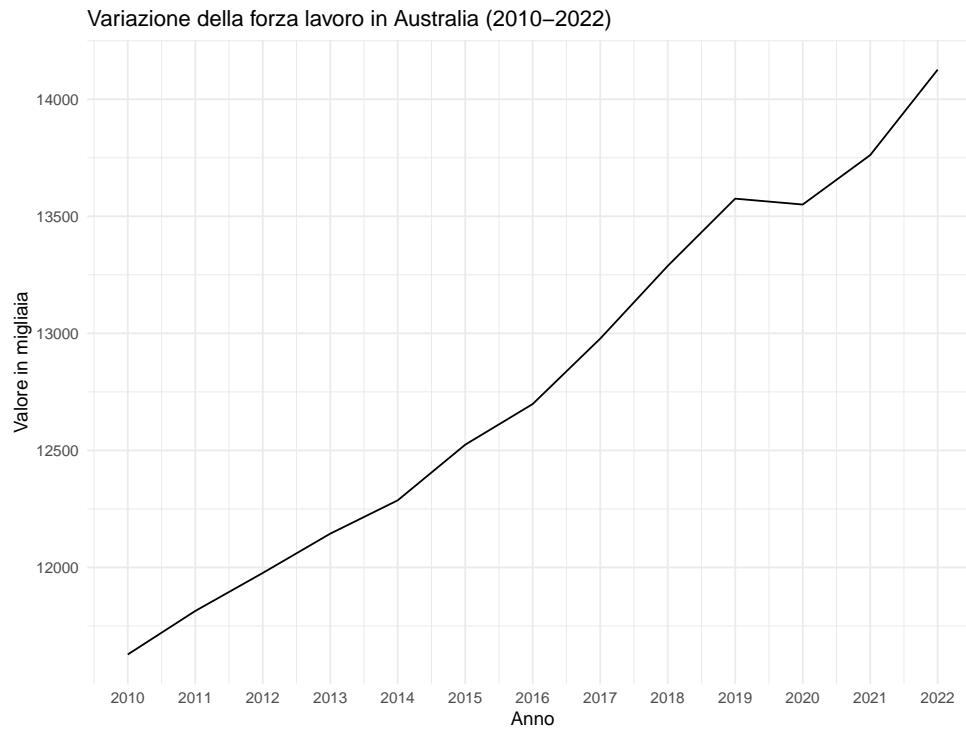
---

### Serie temporali e Scatterplot

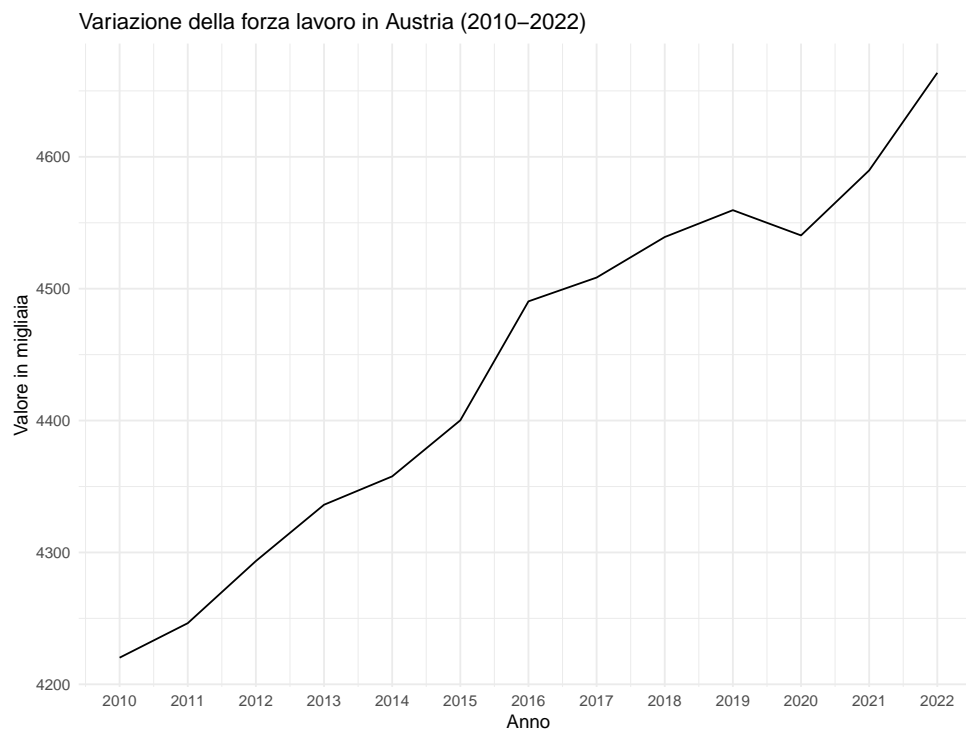
---

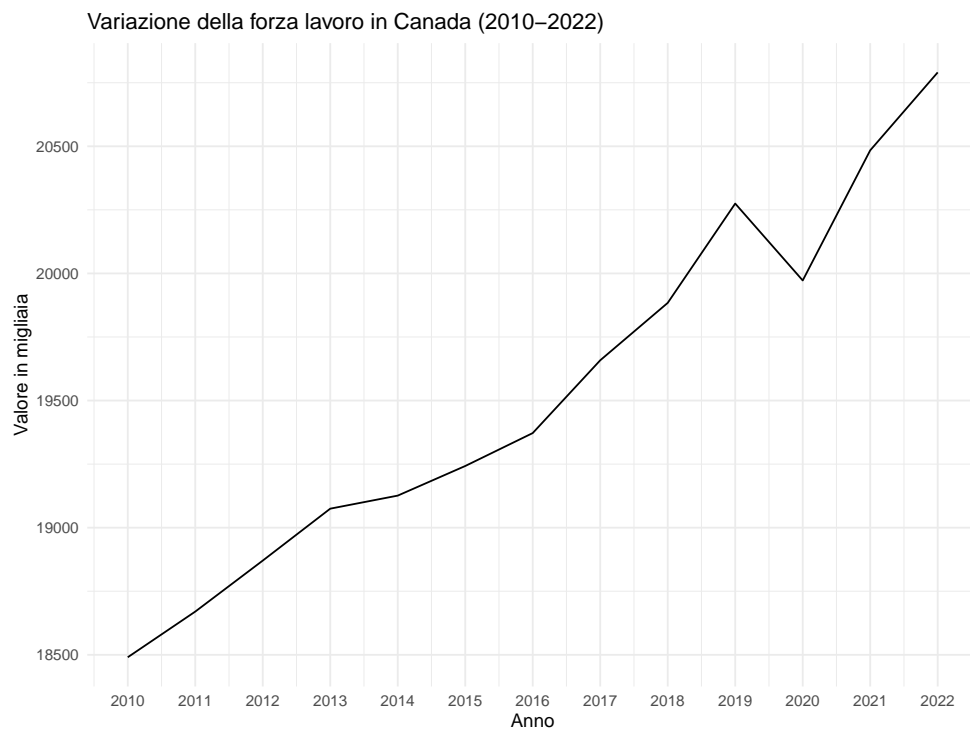
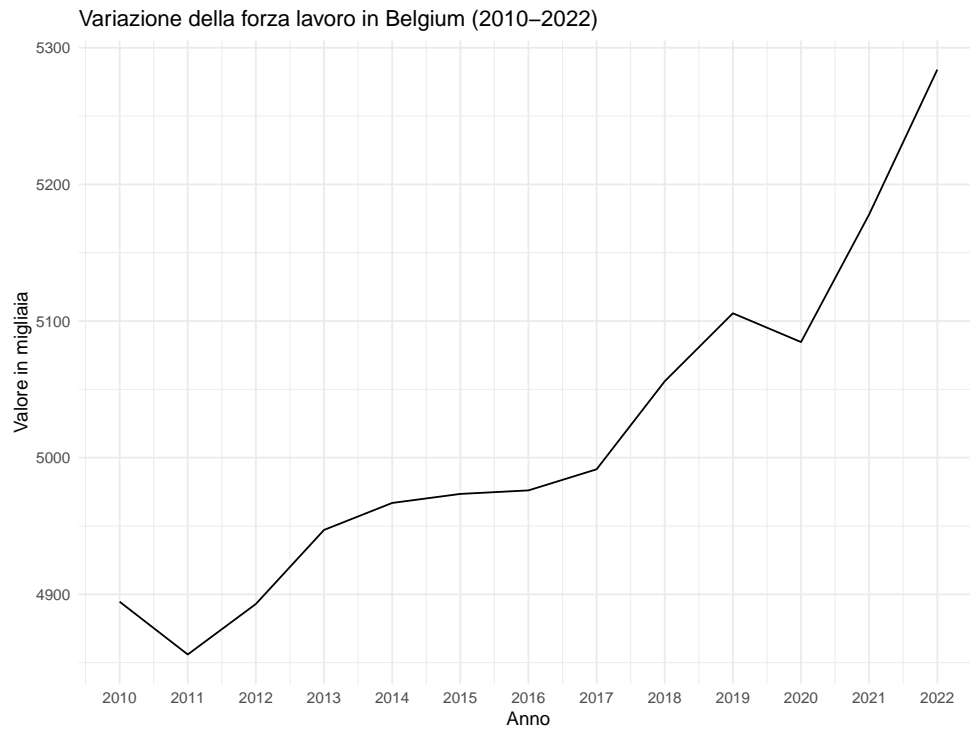
#### 2.1 Serie Temporali

La forza lavoro rappresenta una caratteristica importante per l'economia e la crescita di ogni paese. Di seguito, sono riportate le serie temporali per 38 paesi, dal 2010 al 2022, sulla forza lavoro. È incluso anche "OECD Countries" che contiene la somma dei valori di tutti e 38 i paesi. I grafici appaiono in ordine alfabetico del nome del paese. I grafici più interessanti e/o significativi sono stati commentati.



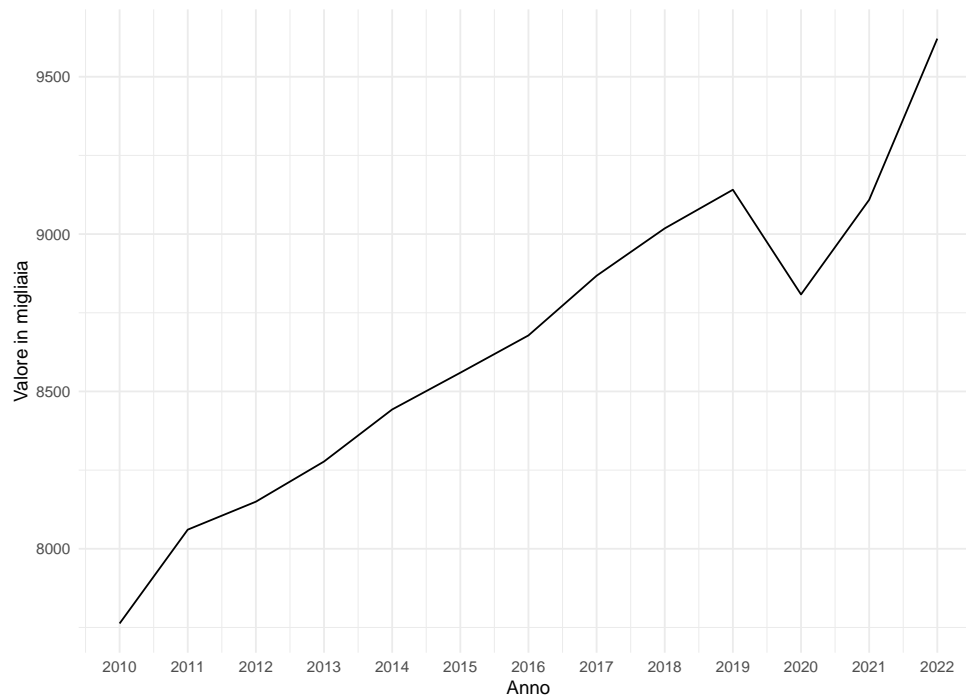
È interessante notare come nel 2020 ci sia stata una diminuzione della forza lavoro in quasi tutti i paesi a causa della pandemia causata dal COVID-19. [1]



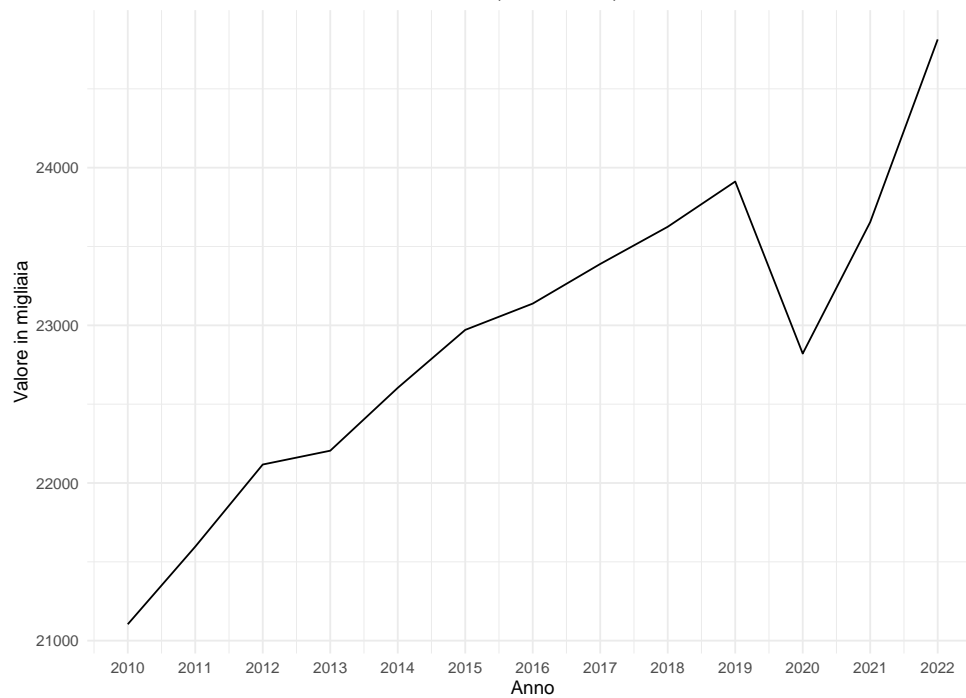


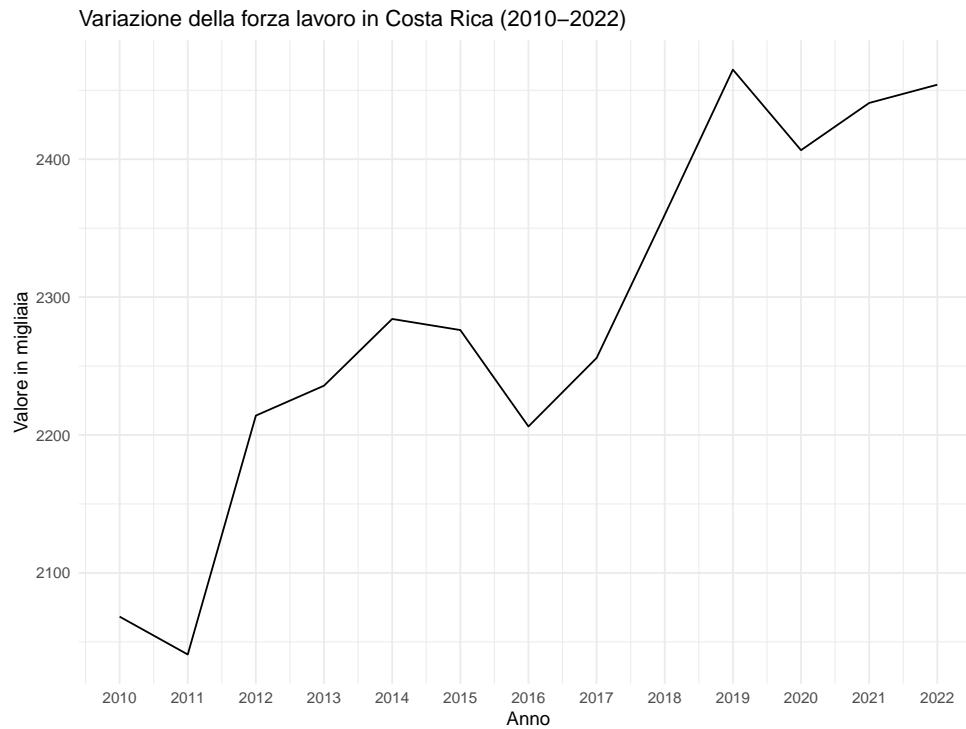


Variazione della forza lavoro in Chile (2010–2022)

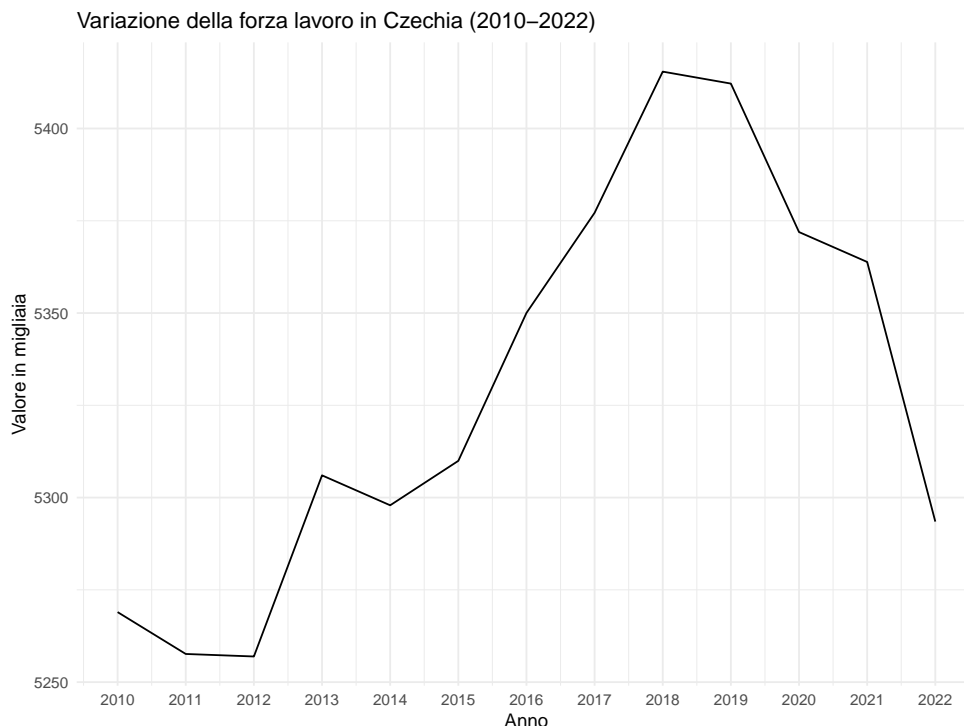


Variazione della forza lavoro in Colombia (2010–2022)





In Costa Rica, nel 2016 c'è stato un calo drastico della forza lavoro. Nel periodo 2008-2015, il rapporto debito pubblico/GDP è aumentato fino al 42.2%. Inoltre, la Costa Rica contiene problemi strutturali da molto tempo: La disuguaglianza economica (disparità economica e sociale tra individui all'interno di un dato paese) ha continuato ad aumentare dalla metà degli anni '90 fino a livelli molto alti; La produttività del lavoro (l'efficienza con cui viene utilizzata la forza lavoro per produrre beni o servizi) è stata deludente a causa del lento progresso tecnologico. [2]



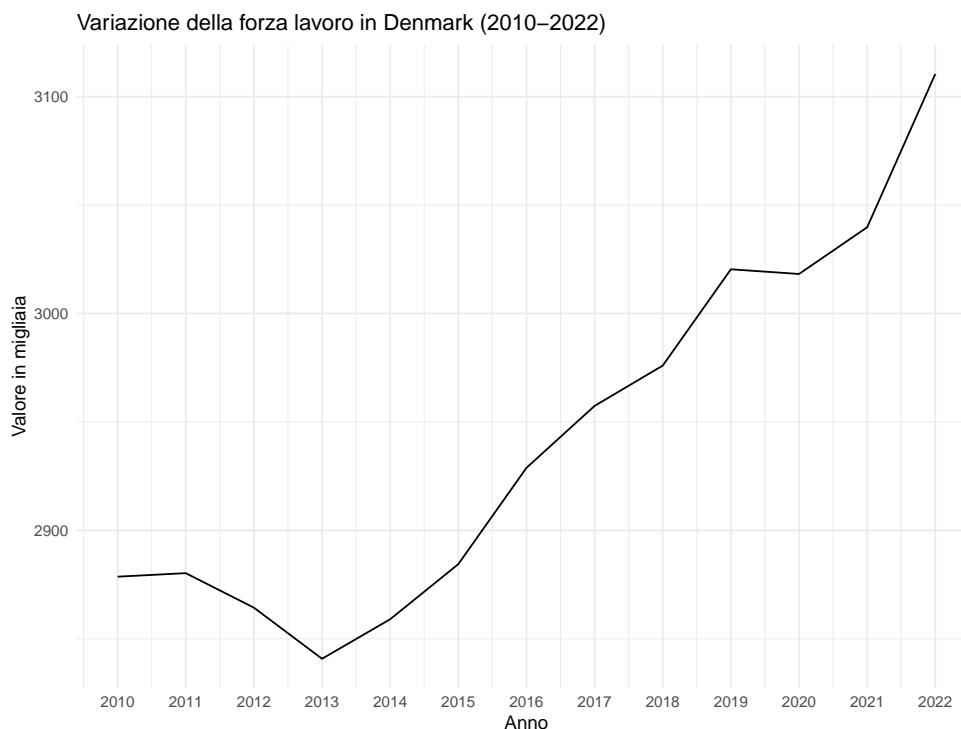
Nel 2018, l'economia ceca era tutto sommato stabile, ben gestita e con un solido equilibrio interno ed esterno. Tuttavia, l'economia sembrava essere un po' surriscaldata, ovvero l'eccessiva offerta di moneta aveva portato ad un aumento della domanda che la produzione non riusciva a reggere, e quindi il PIL era il 2-3% circa rispetto al prodotto potenziale. Alla fine dello scorso anno, nel 2022, l'economia ceca presentava grandi squilibri macroeconomici e un'inflazione molto elevata. I fenomeni mondiali che si sono susseguiti in questi 5 anni, quali la pandemia da COVID-19 (2020), la rottura delle catene di approvvigionamento (2021) e la crisi energetica iniziata lo scorso anno (2022) hanno messo a dura prova l'economia ceca, molto sensibile per conformazione. Si tratta di un'economia piccola e molto aperta in quanto scarsa dal punto di vista delle fonti energetiche e dipendente da catene di approvvigionamento complesse.

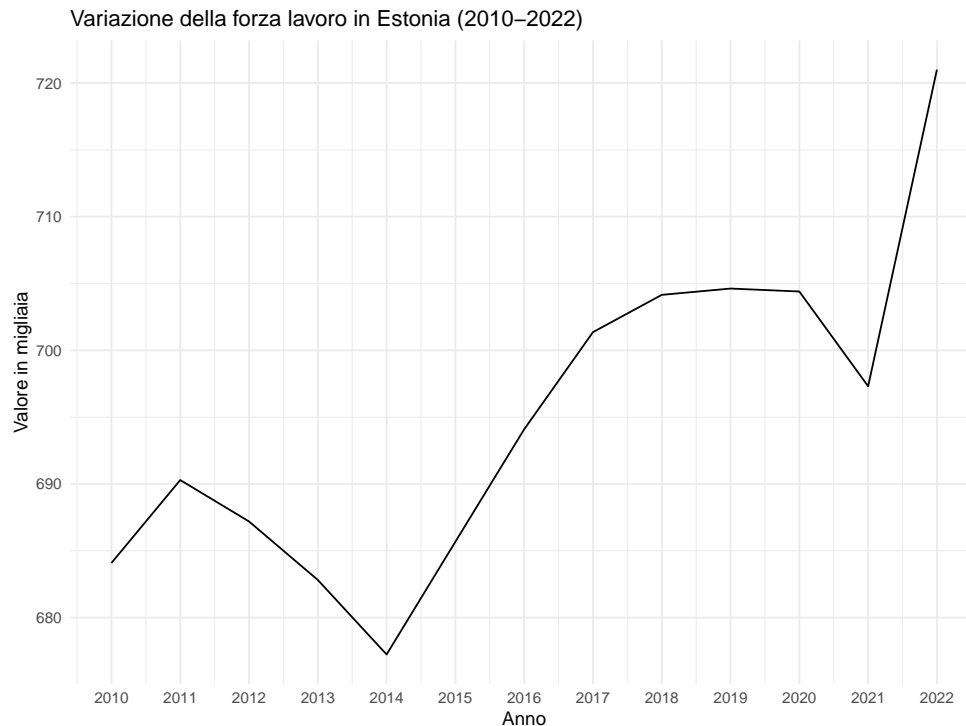
La crisi energetica ha colpito duramente la Repubblica Ceca, che in quanto Paese piccolo ha scarsa diversità energetica ed è costretta ad importare la maggior parte del gas naturale dalla Russia, dipendendo dai suoi prezzi, che sono in aumento. L'aumento ha poi causato un effetto inflazionistico dei prezzi al consumo, impattando negativamente la situazione economica delle famiglie e del bilancio statale.

A fronte dei problemi globali che l'economia si è trovata ad affrontare, le politiche

economiche non sono state delle migliori: il populismo nella politica fiscale pre-pandemia e i cambiamenti nella tassazione hanno impattato negativamente le entrate di bilancio (significativo divario del bilancio pubblico) e la crescita del PIL. Sono state valorizzate eccessivamente le pensioni, è stata abolita la tassa sul salario lordo, riducendo i gettiti al bilancio statale. Oltretutto, sono stati concessi sostegni fiscali alle famiglie, misure costose e non mirate, che hanno aumentato il debito pubblico. Anche gli interventi macroeconomici nel mercato immobiliare sono stati tardivi, causando un ulteriore aumento dell'inflazione.

I problemi di inflazione, divario commerciale, peggioramento del bilancio pubblico, sono iniziati prima del 2018, quando nel 2017 l'economia ha cominciato a surriscaldarsi. Infatti, il saldo del bilancio del governo era in deficit già nel 2019, prima della pandemia. Per limitare i danni, le politiche fiscali e monetarie avrebbero dovuto essere più restrittive. La pandemia chiaramente ha peggiorato la situazione, è stato un periodo di "stallo" che ha indebolito l'economia. Anche il mercato del lavoro è in crisi: la mancanza di forza lavoro disponibile è un limite per la crescita del PIL, e le pressioni sui salari e sui costi sono aumentate contribuendo ad una crescente inflazione. [3]

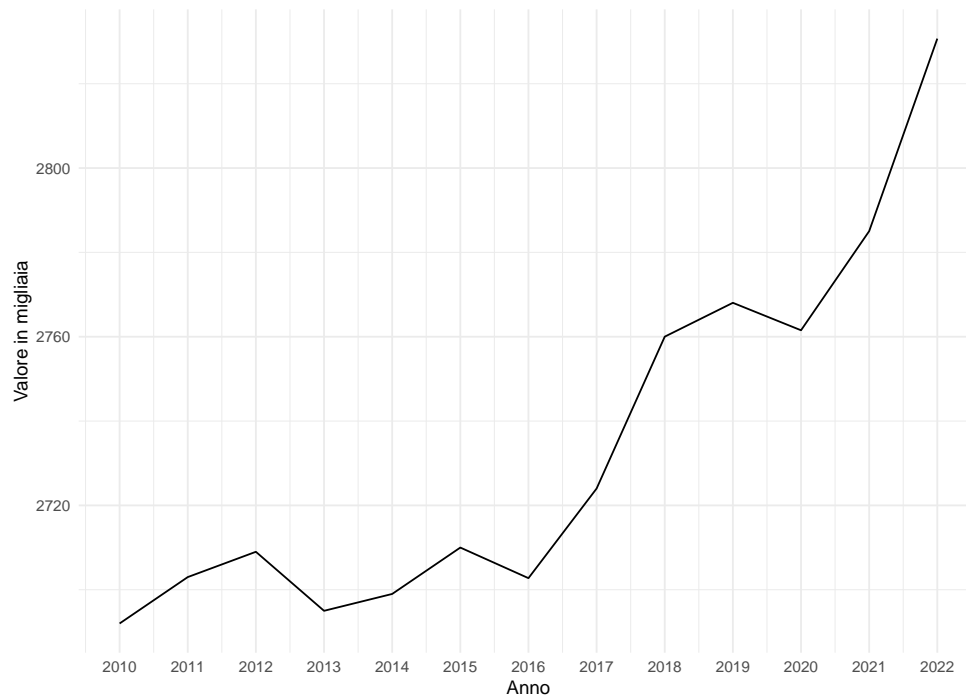




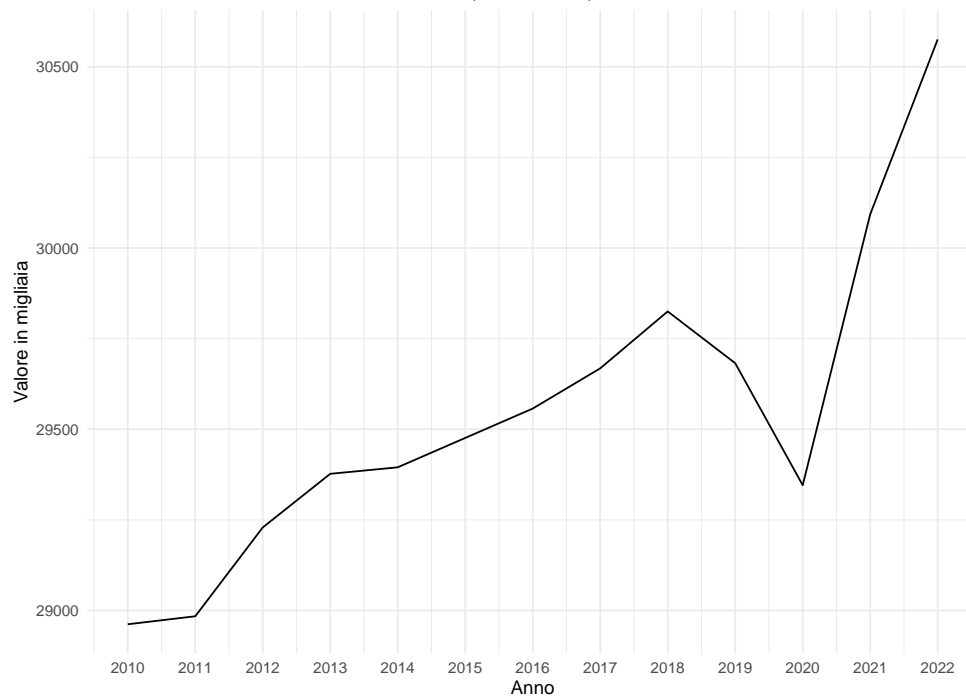
Durante il 2014, l'Estonia ha sperimentato un significativo aumento dell'emigrazione verso altri paesi europei, in particolare verso i paesi dell'Unione Europea con mercati del lavoro più forti. Questo potrebbe aver ridotto il numero di persone disponibili nella forza lavoro dell'Estonia.

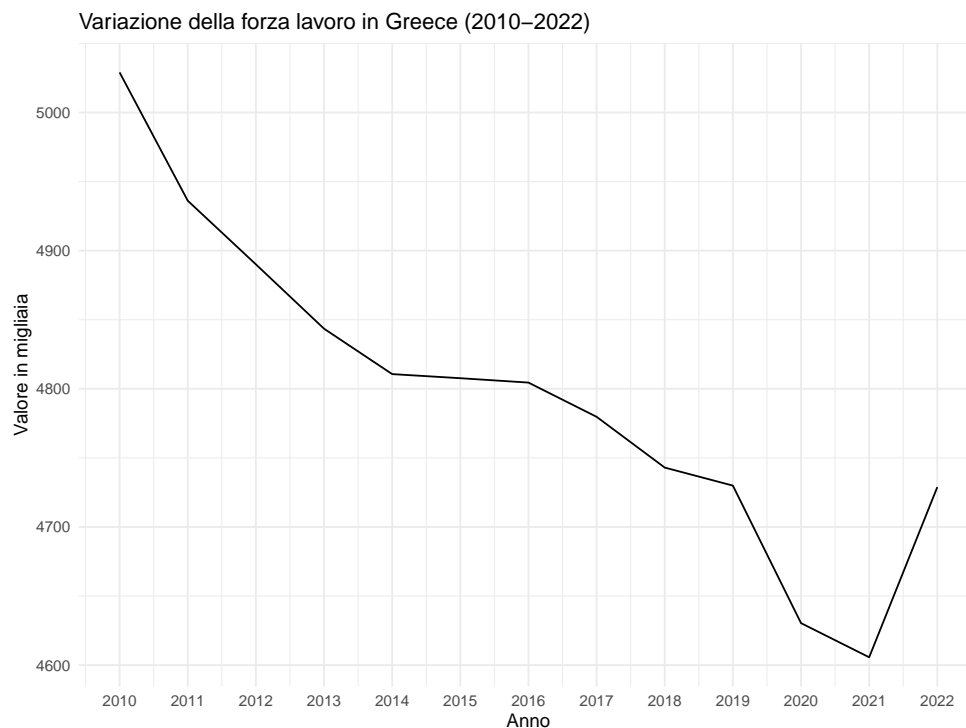
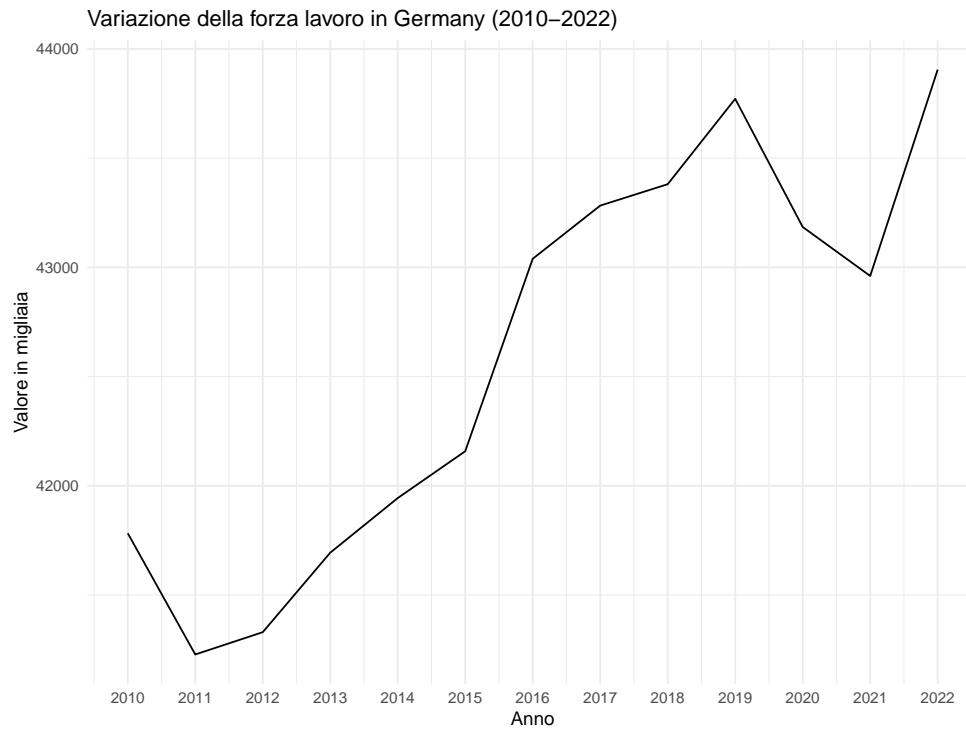
Anche se l'Estonia non ha subito gli stessi livelli di impatto della crisi economica globale del 2008 come alcuni altri paesi europei, potrebbero essere state presenti ancora delle conseguenze nel 2014. La crisi economica potrebbe aver portato a licenziamenti, chiusure di aziende e riduzione delle opportunità di lavoro, riducendo di conseguenza la forza lavoro.

Variazione della forza lavoro in Finland (2010–2022)



Variazione della forza lavoro in France (2010–2022)





La Grecia ha vissuto una crisi economica molto grave già a partire dal 2010 (osservando il grafico), la quale ha avuto un impatto significativo sulla forza lavoro nel paese.

La crisi finanziaria del 2008 ha colpito pesantemente l'economia greca, aggravata

da un alto livello di debito pubblico accumulato nel corso degli anni precedenti. L'aumento del debito pubblico ha portato a una crisi di fiducia nei confronti dell'economia greca, causando un calo degli investimenti esteri e una stretta del credito.

Per ottenere l'assistenza finanziaria internazionale dall'Unione Europea (UE), dalla Banca Centrale Europea (BCE) e dal Fondo Monetario Internazionale (FMI) - noti collettivamente come "troika" - la Grecia ha dovuto implementare severe misure di austerità economica. Queste misure includevano tagli alla spesa pubblica, aumenti delle tasse e riforme strutturali, che hanno portato a una riduzione dei posti di lavoro nel settore pubblico e privato.

La crisi ha portato anche al collasso di diverse banche greche, con conseguente perdita di risparmi dei cittadini e un impatto negativo sul sistema finanziario del paese. Il crollo delle banche ha ulteriormente indebolito l'economia greca e ha contribuito alla fuga di capitali e all'aumento della disoccupazione.

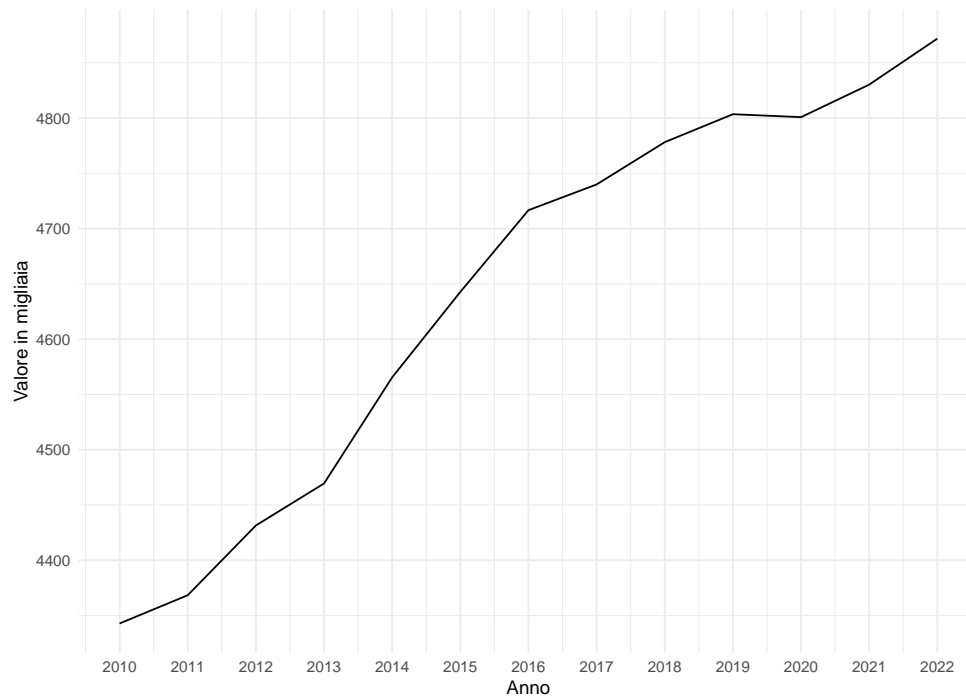
A causa della crisi finanziaria e delle misure di austerità, gli investimenti e la fiducia degli investitori sono diminuiti, il che ha causato una riduzione della crescita economica e un calo delle opportunità di lavoro.

Molte persone in cerca di lavoro in Grecia hanno cercato opportunità all'estero a causa della crisi economica, portando a un esodo di lavoratori qualificati e giovani, il che ha ulteriormente aggravato la situazione della forza lavoro nel paese.

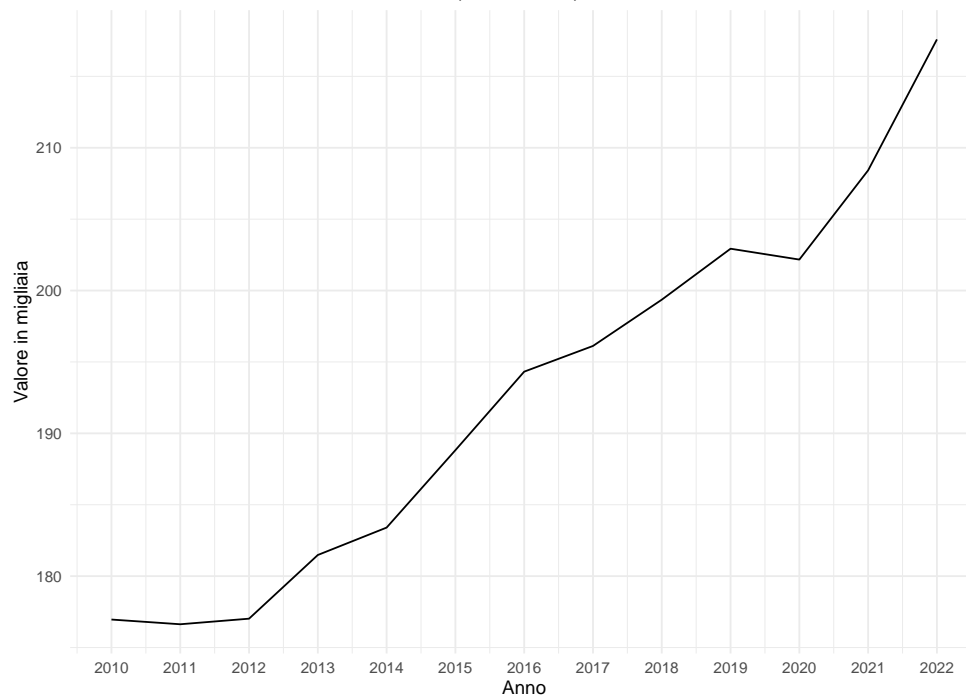
Nel tempo, nonostante gli sforzi per implementare riforme economiche e strutturali, la Grecia ha faticato a riprendersi completamente dalla crisi. La combinazione di una serie di fattori come l'austerità, il debito elevato, la mancanza di investimenti e la fuga di capitali ha ostacolato la ripresa economica e il miglioramento delle condizioni della forza lavoro nel paese.[4]



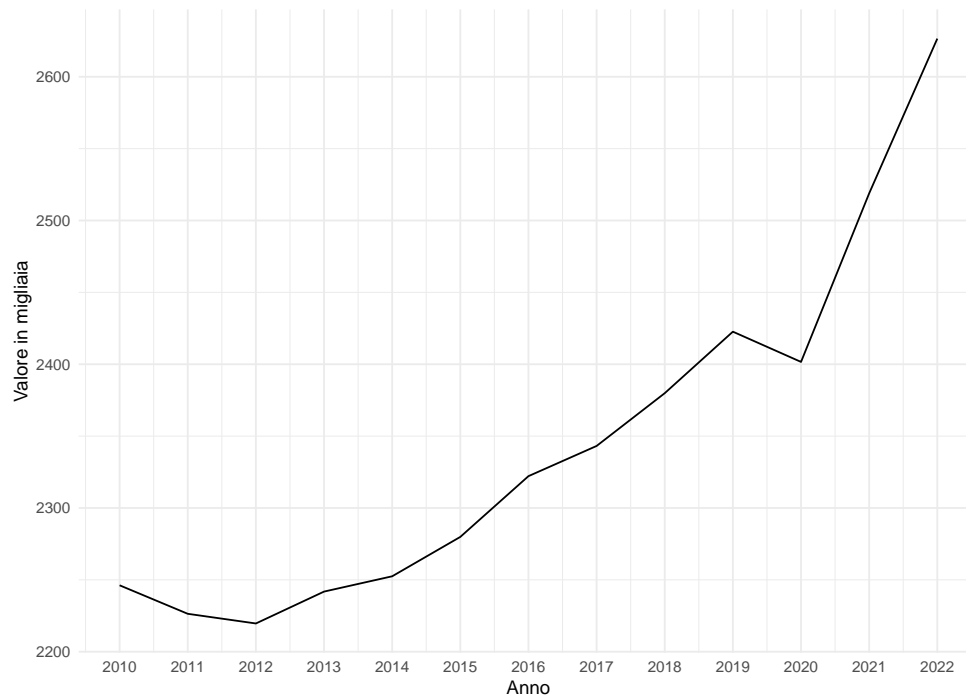
Variazione della forza lavoro in Hungary (2010–2022)



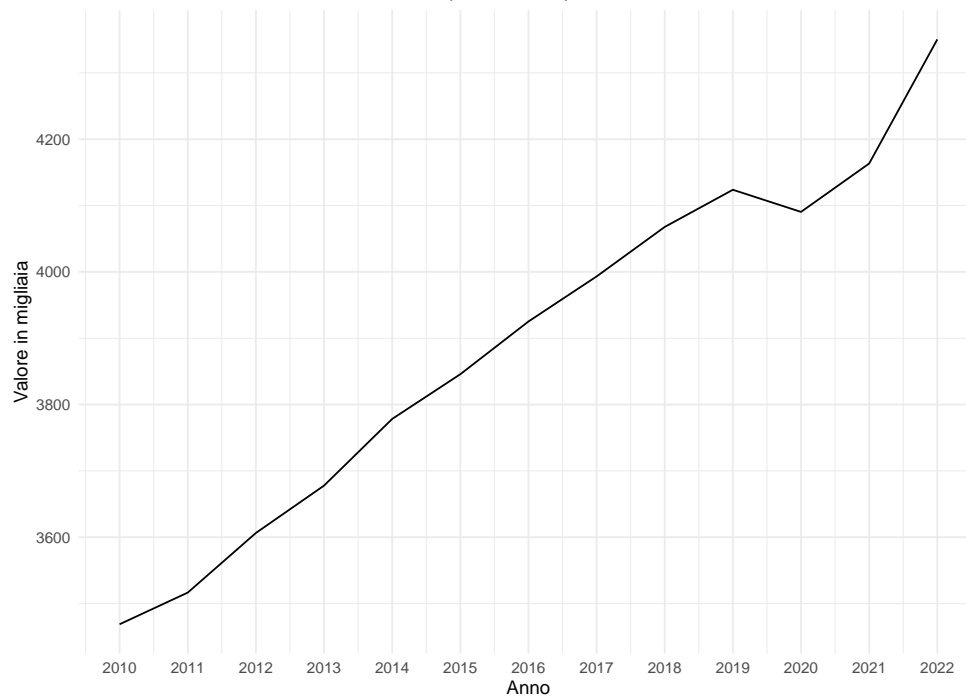
Variazione della forza lavoro in Iceland (2010–2022)

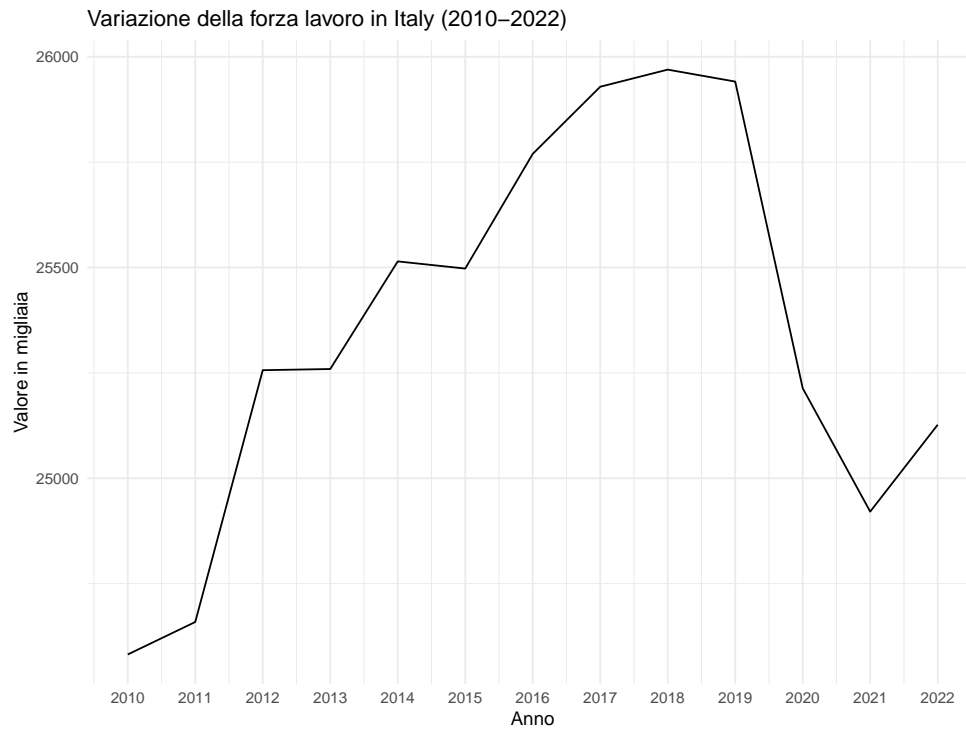


Variazione della forza lavoro in Ireland (2010–2022)

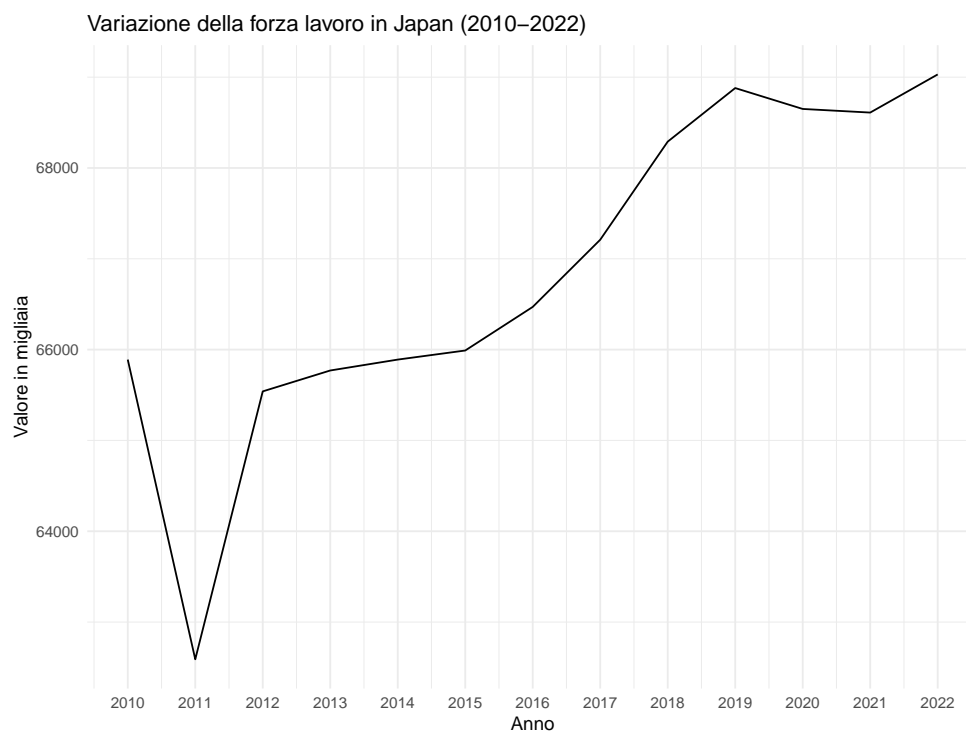


Variazione della forza lavoro in Israel (2010–2022)





L'Italia è stata la nazione che ha subito più di tutte la crisi legata al COVID-19, infatti nel grafico è possibile notare il calo della forza lavoro nel 2020 e nel 2021. È probabile che ci sia stata una diminuzione della forza lavoro a causa dell'aumento di persone anziane nella popolazione, come vedremo più avanti.

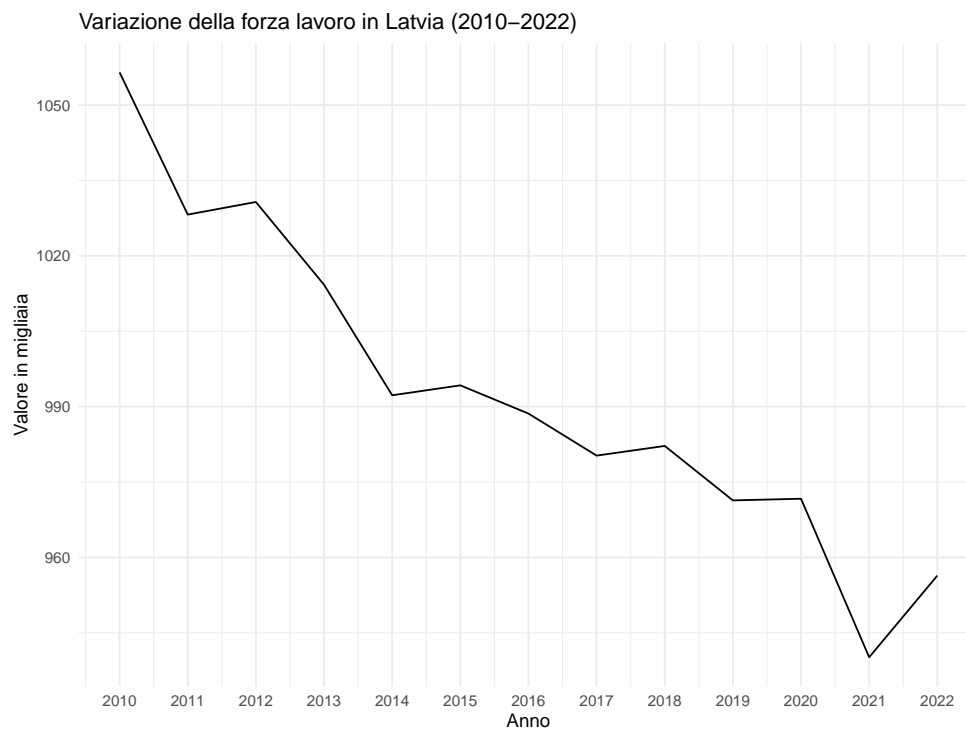
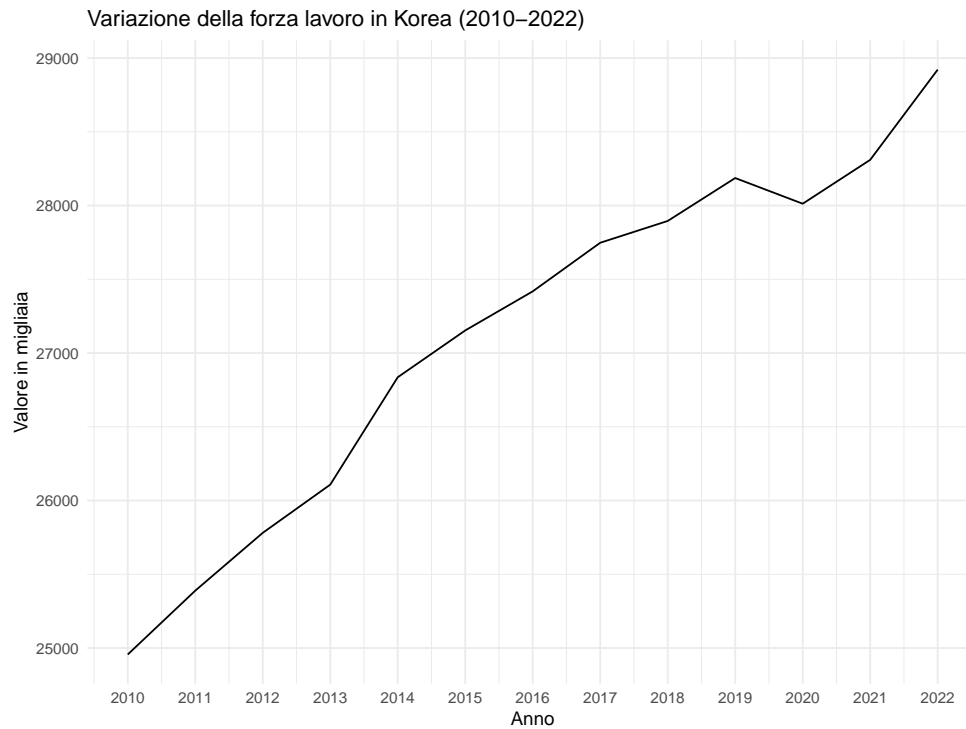


Il brusco calo della forza lavoro in Giappone nel 2011 è stato principalmente influenzato da due fattori principali;

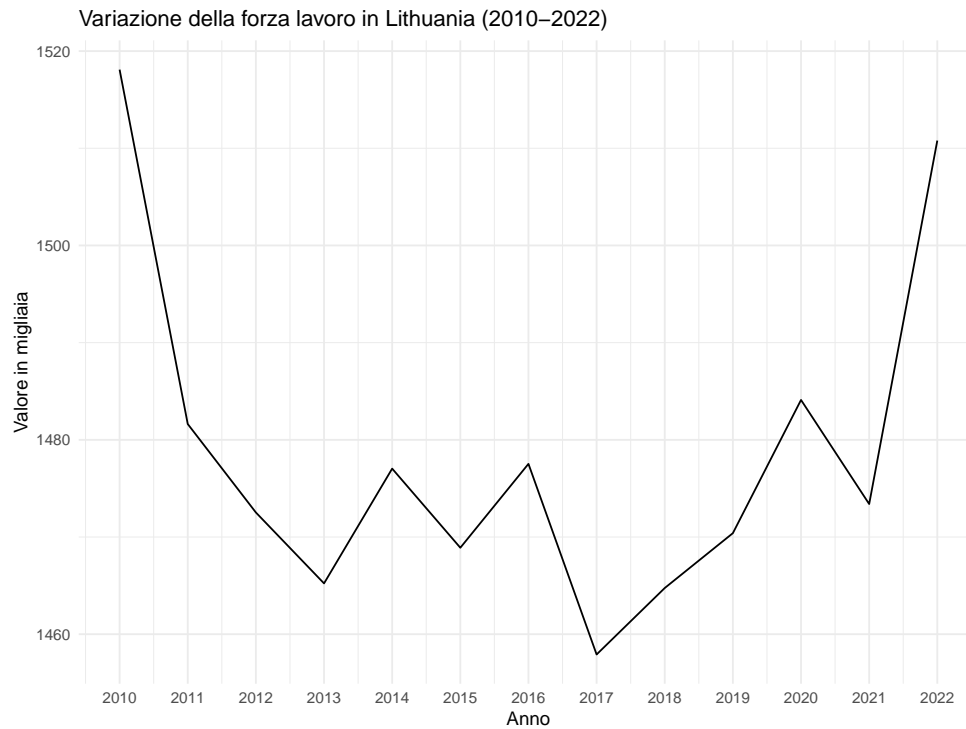
Disastri naturali: Nel marzo 2011, il Giappone è stato colpito da un grave terremoto seguito da uno tsunami devastante nella regione di Tohoku [5]. Questi eventi naturali hanno causato ingenti danni materiali, interruzioni delle infrastrutture e perdite umane significative. Le conseguenze di questi disastri naturali hanno influenzato negativamente l'economia del paese, riducendo temporaneamente l'attività economica e causando un calo della forza lavoro in settori come la costruzione, il trasporto e il turismo.

Incidenti nucleari di Fukushima: Il terremoto e lo tsunami del marzo 2011 hanno causato gravi danni alla centrale nucleare di Fukushima Daiichi, causando una serie di incidenti nucleari e la contaminazione radioattiva nelle zone circostanti. A seguito di questi incidenti, molte persone hanno abbandonato le loro case e i loro posti di lavoro nelle aree colpite, portando a una significativa riduzione della forza lavoro in quelle regioni. [6]

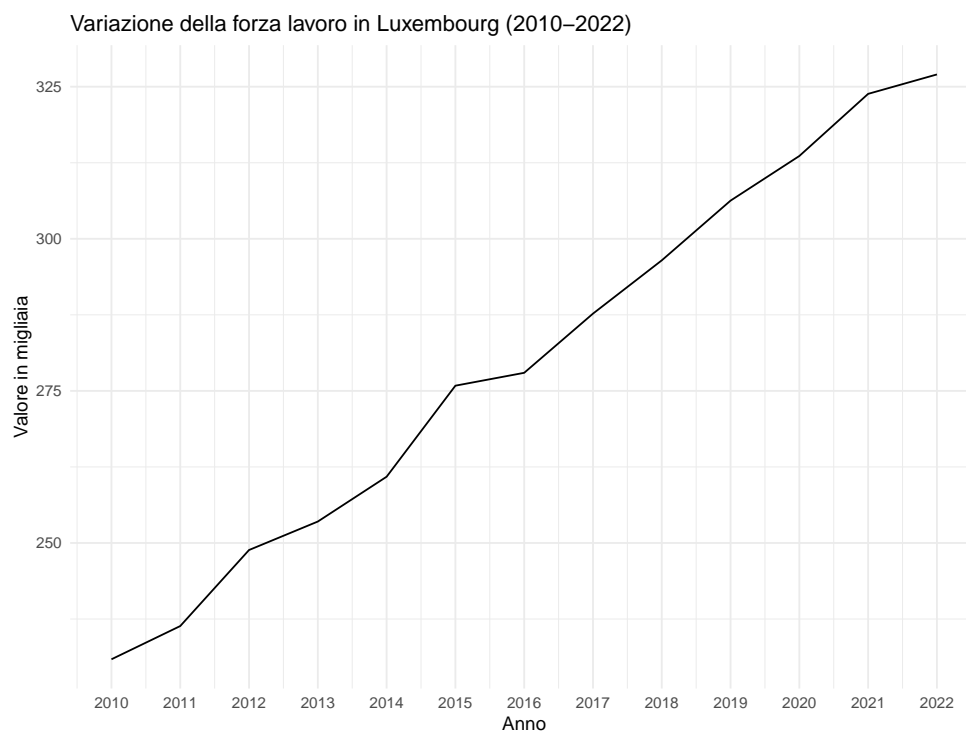
Questi eventi hanno avuto un impatto immediato sulla forza lavoro in Giappone nel 2011, con una diminuzione temporanea del numero di lavoratori impiegati in settori specifici e nelle regioni colpite dai disastri naturali e nucleari. Tuttavia, è importante notare che il Giappone ha rapidamente avviato sforzi di ricostruzione e recupero per affrontare gli effetti di questi eventi e ripristinare l'attività economica e la forza lavoro nelle aree colpite.



La Lettonia ha avuto un'evoluzione sulla forza lavoro molto simile a quella della Grecia. Infatti entrambi i paesi hanno avuto molte difficoltà a reagire alla crisi del 2008. Inoltre, la Lettonia soffre anche di un forte problema di emigrazione. [7]



Anche la Lituania, come la Grecia e la Lettonia, è stata fortemente condizionata dalla crisi del 2008. Tuttavia, a differenza della maggior parte dei paesi europei, la Lituania è stata in grado di riprendersi rapidamente negli anni successivi e la produttività è tornata a livelli vicini al periodo pre-crisi. [8]



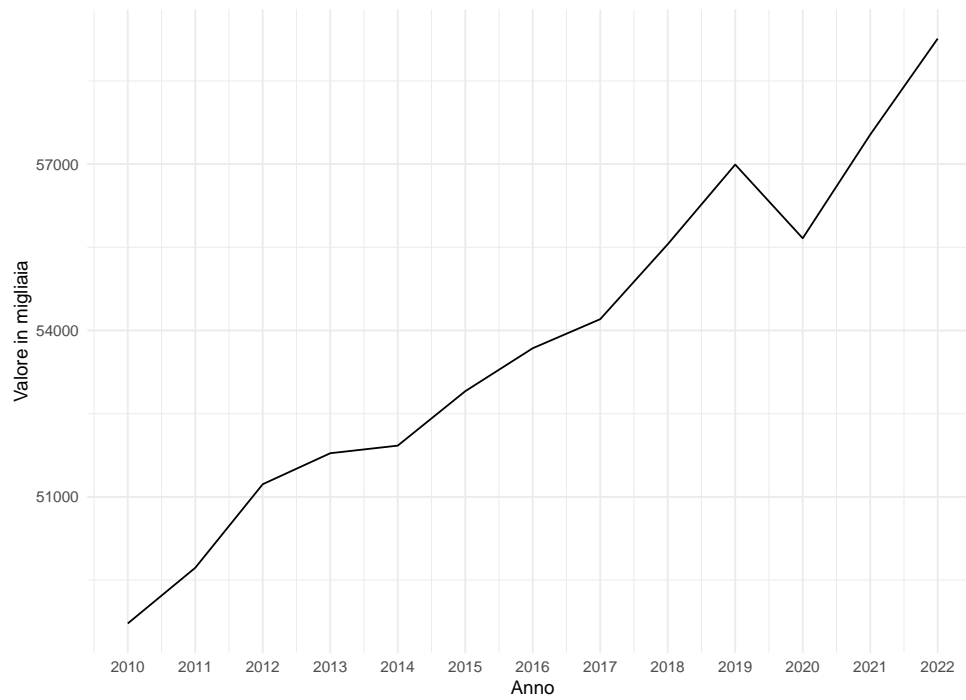
Osservando il grafico, colpisce come non ci sia stato un calo della forza lavoro nel 2021.

Il Lussemburgo è un paese piccolo, multilingue e molto aperto. È il secondo paese più piccolo dell'Unione Europea sia in termini di superficie (2 586  $km^2$ ) che di popolazione (645 397 abitanti al 1 gennaio 2022), dopo Malta. Con 259,4 abitanti per  $km^2$  nel 2021, è tra i cinque paesi con la più alta densità di popolazione nell'UE e nell'OCSE. Il lussemburghese è la lingua ufficiale del paese, ma il paese ha anche il francese e il tedesco come lingue ufficiali amministrative. Inoltre, il Lussemburgo è politicamente ed economicamente ben integrato ed è un membro fondatore dell'UE, dell'eurozona, dell'OCSE, della NATO e dei Paesi Bassi, tra gli altri.

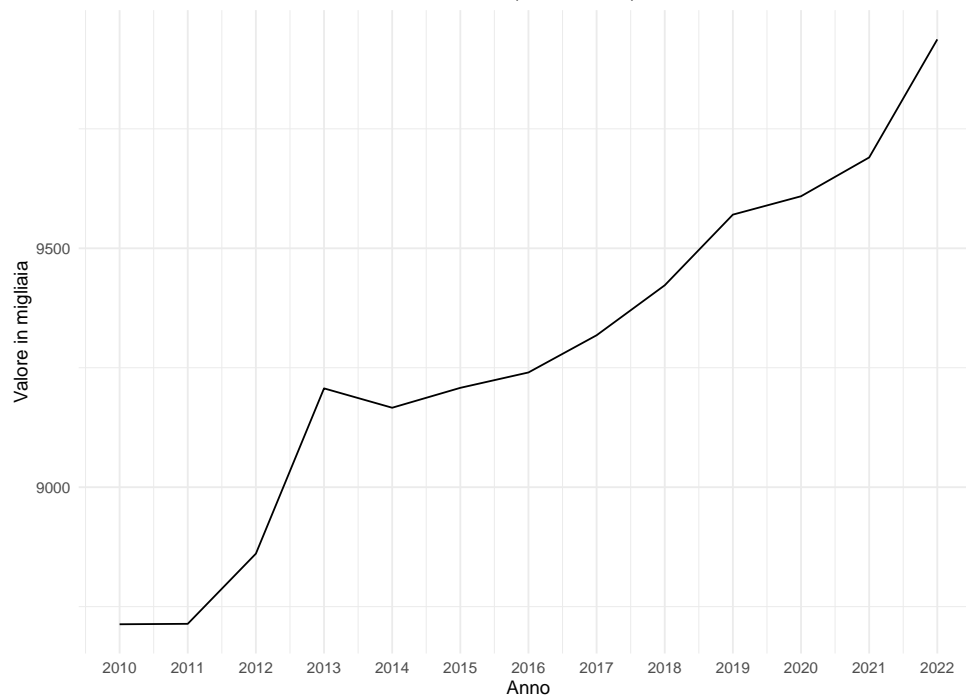
La popolazione del Lussemburgo è in crescita e rimane complessivamente giovane. Il paese ha sperimentato una crescita dinamica della popolazione, con una media del 2% negli ultimi cinque anni, rispetto a una media dell'0,6 per cento in tutti i paesi membri dell'OCSE. Questo forte aumento della popolazione del Lussemburgo è principalmente dovuto a un'alta percentuale di immigrati. Questo afflusso di lavoratori transfrontalieri e immigrati negli ultimi due decenni è anche il motivo per cui la popolazione del Lussemburgo è relativamente giovane; la popolazione del paese è anche proporzionalmente più giovane della media dell'OCSE.

Il sistema politico del Lussemburgo è stabile e gode di un alto livello di fiducia pubblica. Il sistema politico del Lussemburgo è stabile da diversi decenni. Il Lussemburgo è una democrazia parlamentare con una monarchia costituzionale. Il Granduca, un sovrano ereditario, è il capo di stato, ma in pratica il potere esecutivo è esercitato dal governo, guidato dal primo ministro. Durante gli ultimi tre decenni, il governo del Lussemburgo, operando sulla base di maggioranze parlamentari, ha dimostrato in gran parte continuità e stabilità politica. [9]

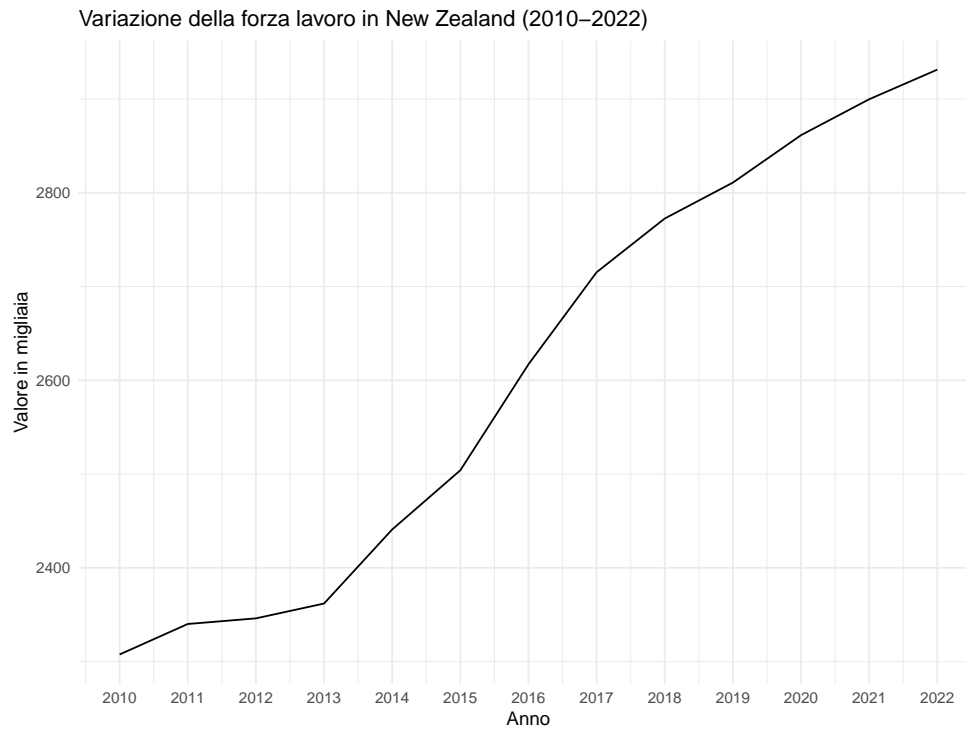
Variazione della forza lavoro in Mexico (2010–2022)



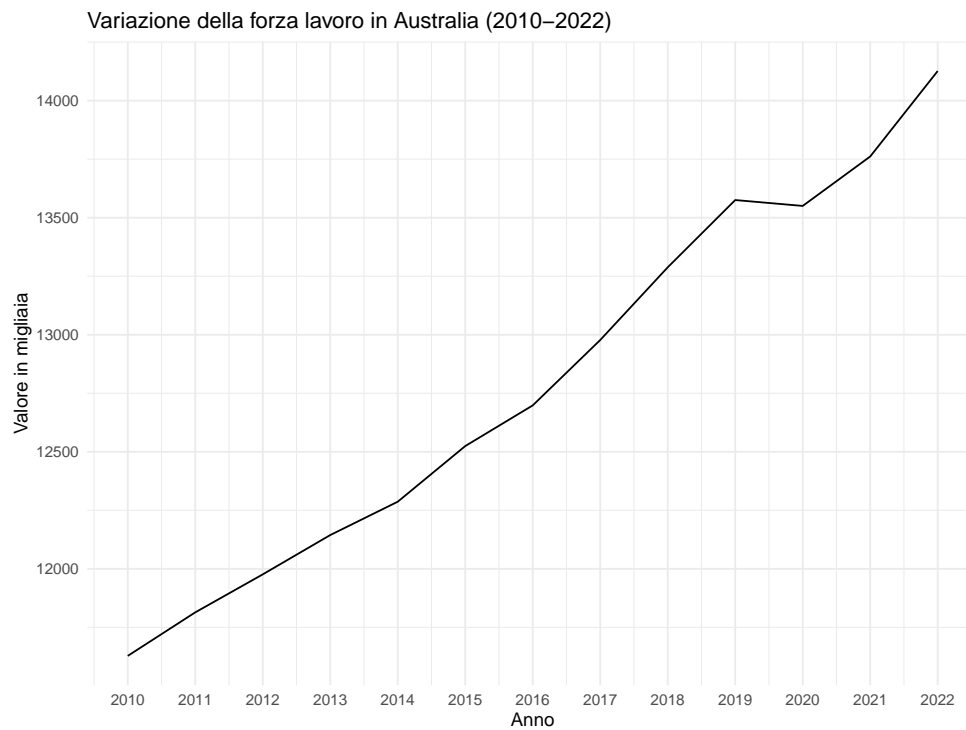
Variazione della forza lavoro in Netherlands (2010–2022)

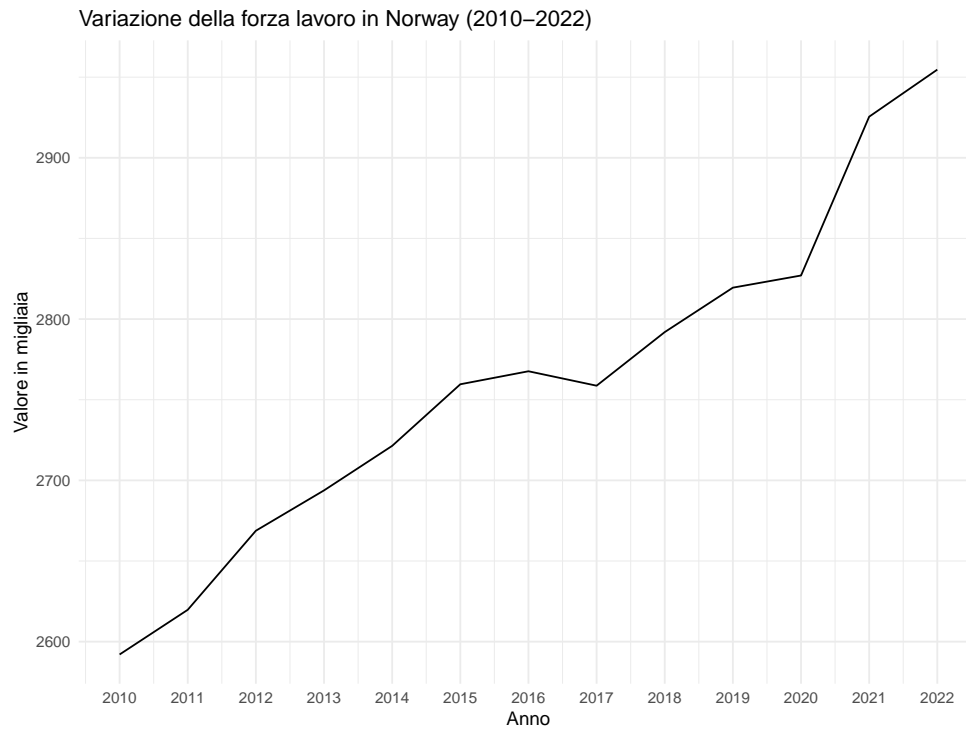




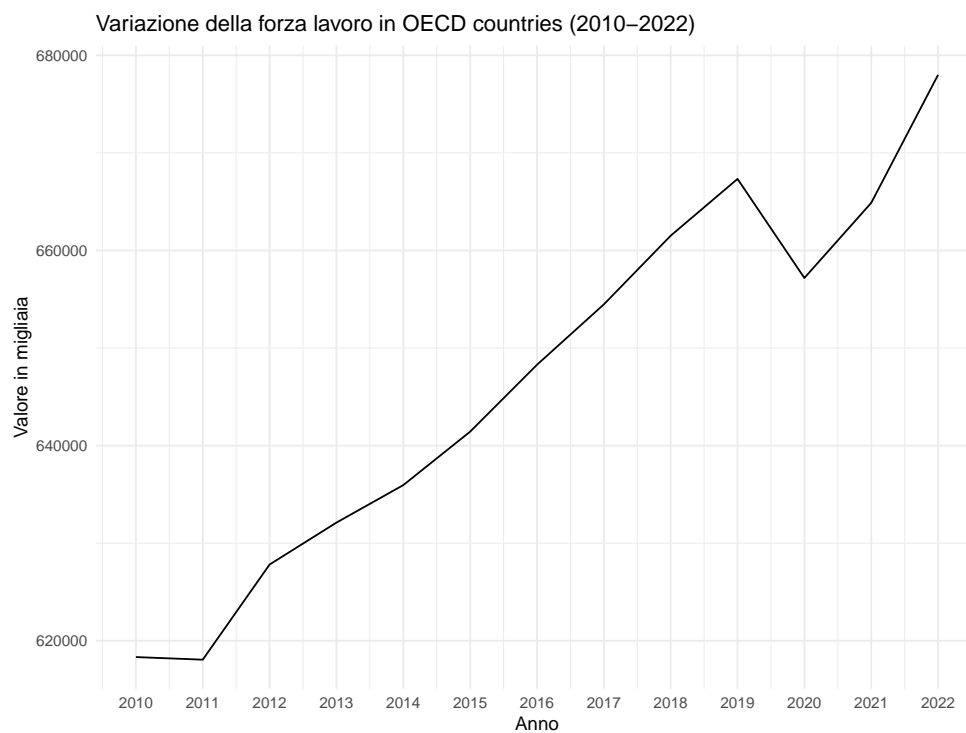


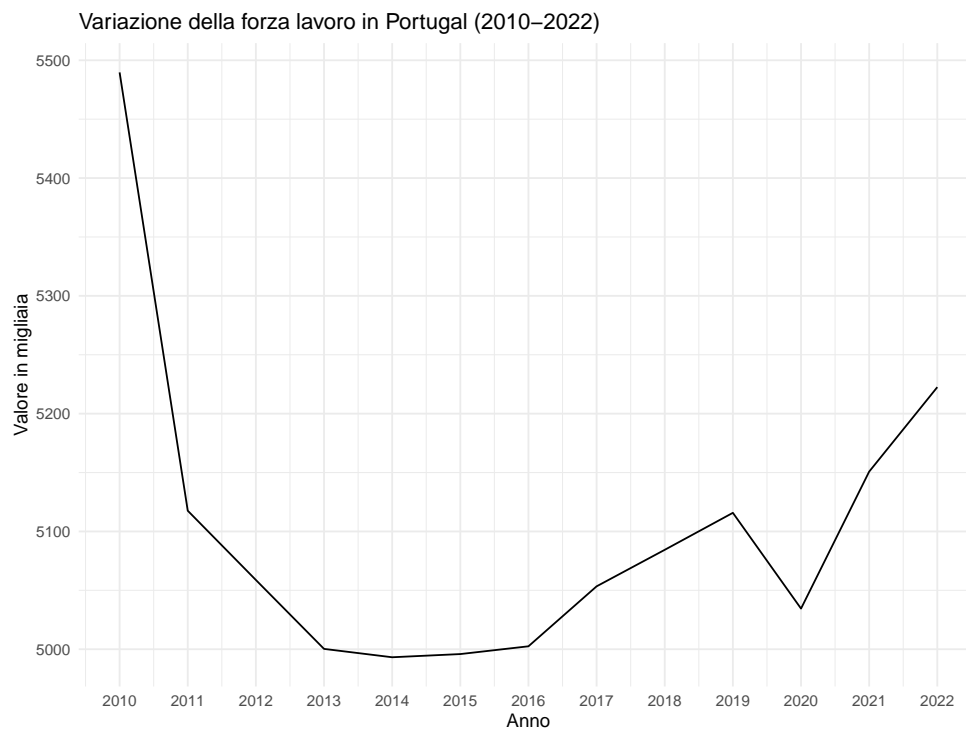
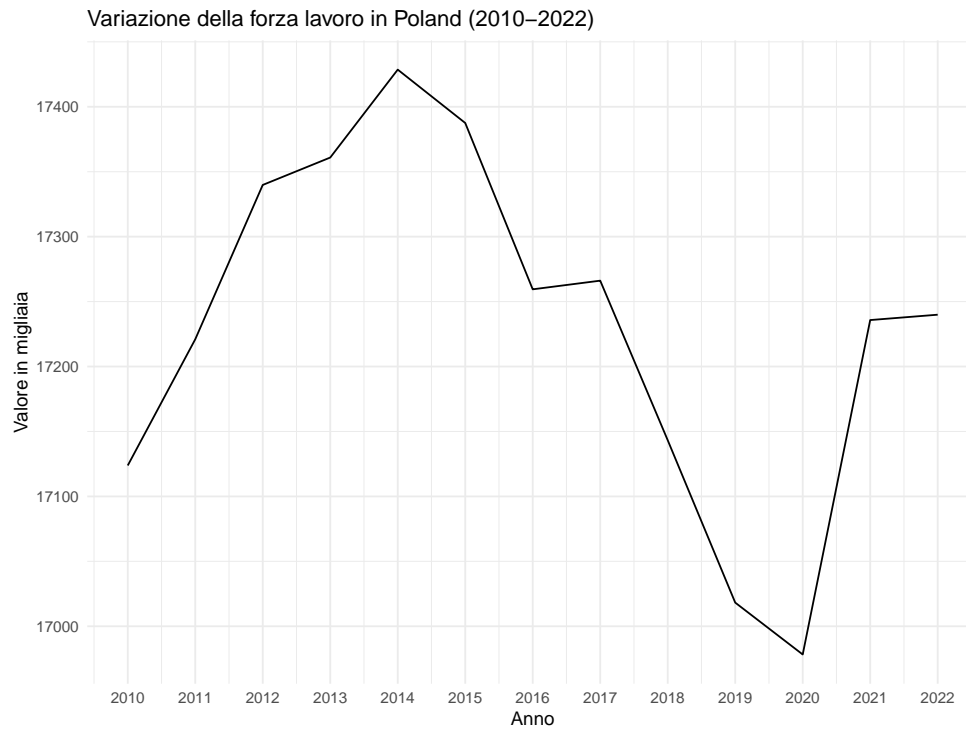
Come il Lussemburgo, anche la Nuova Zelanda ha reagito bene alla pandemia.





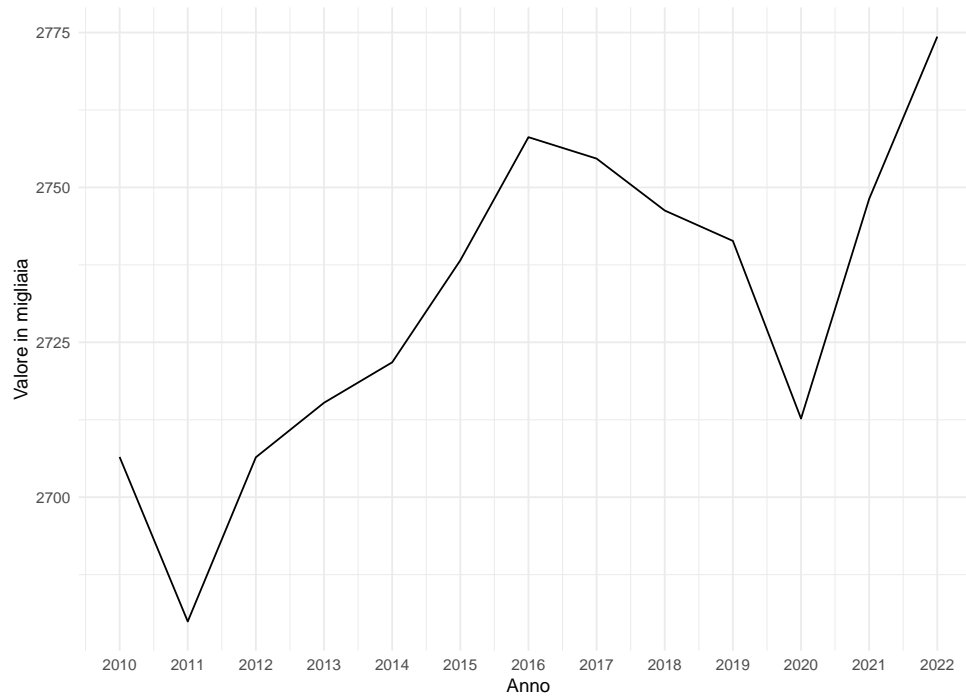
La Norvegia è il paese che meglio ha reagito al COVID-19, come è possibile notare dall'aumento della forza lavoro tra il 2020 al 2021.





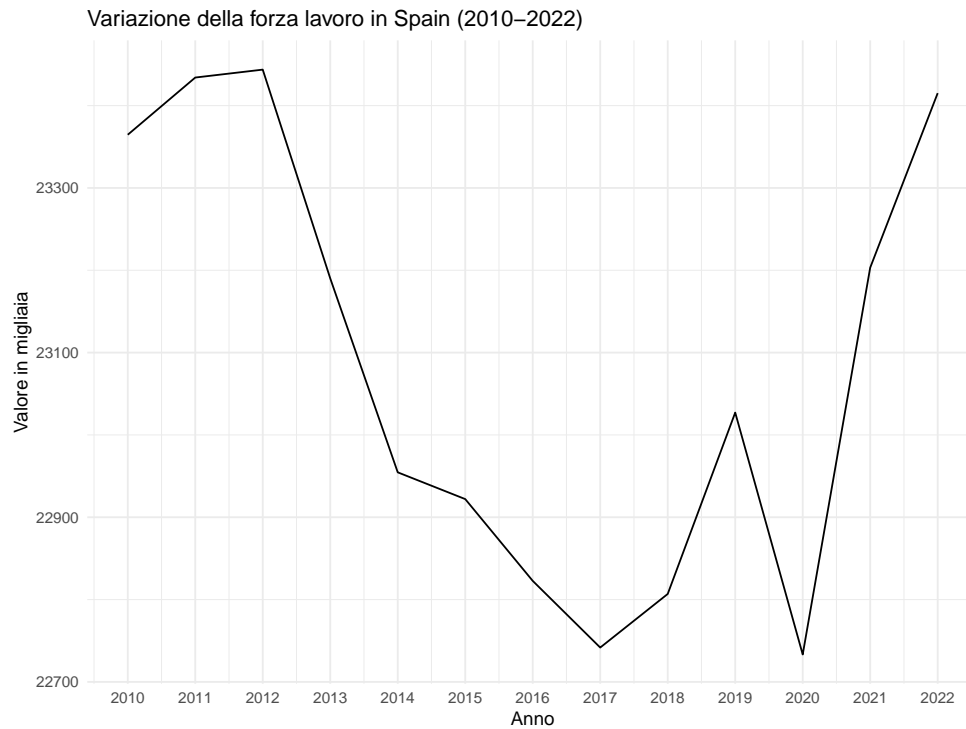
Per quanto riguarda il Portogallo, come abbiamo già analizzato in precedenza, la crisi finanziaria del 2008 ha causato una diminuzione degli investimenti, una contrazione della domanda interna e un aumento della disoccupazione.

Variazione della forza lavoro in Slovak Republic (2010–2022)

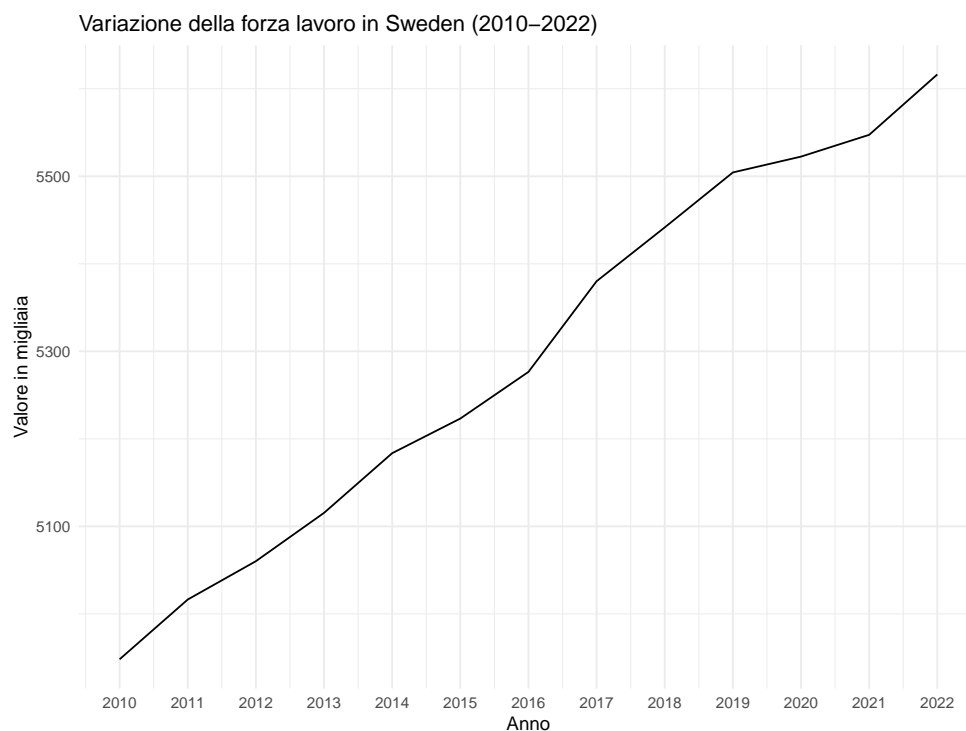


Variazione della forza lavoro in Slovenia (2010–2022)

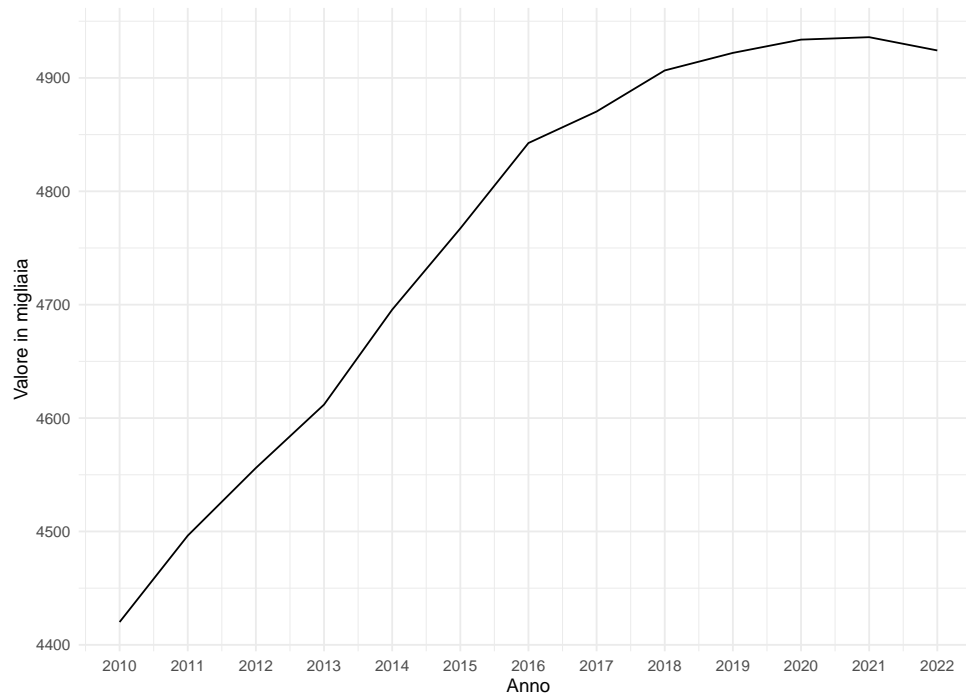




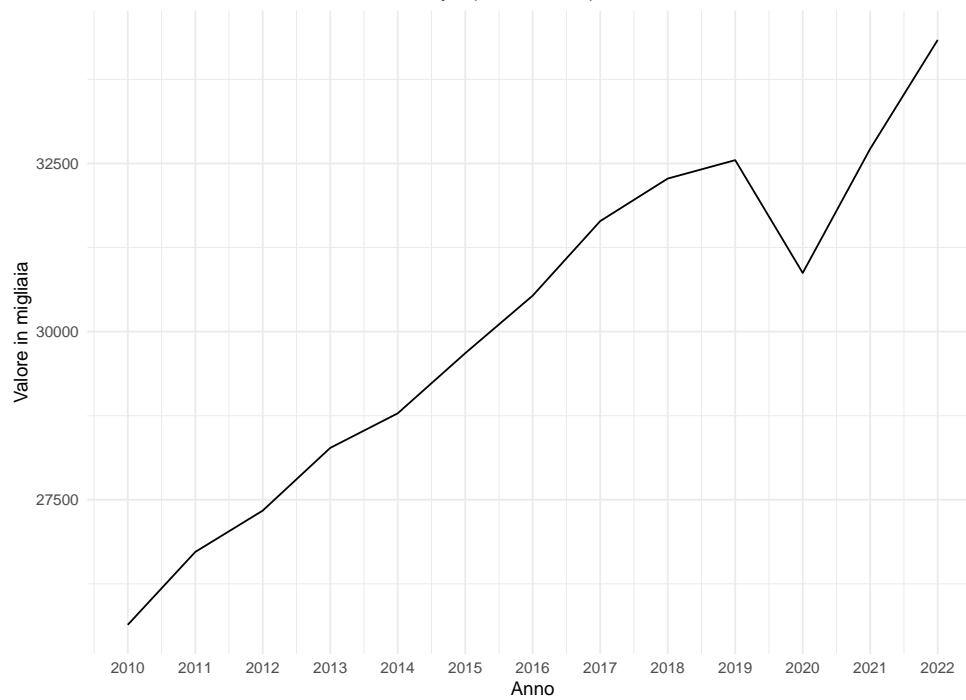
La Spagna è stata fortemente influenzata dalla crisi finanziaria del 2008. Inoltre, il settore immobiliare in Spagna è stato particolarmente colpito da una bolla immobiliare scoppiata, che ha portato a una diminuzione degli investimenti nel settore e alla perdita di posti di lavoro.

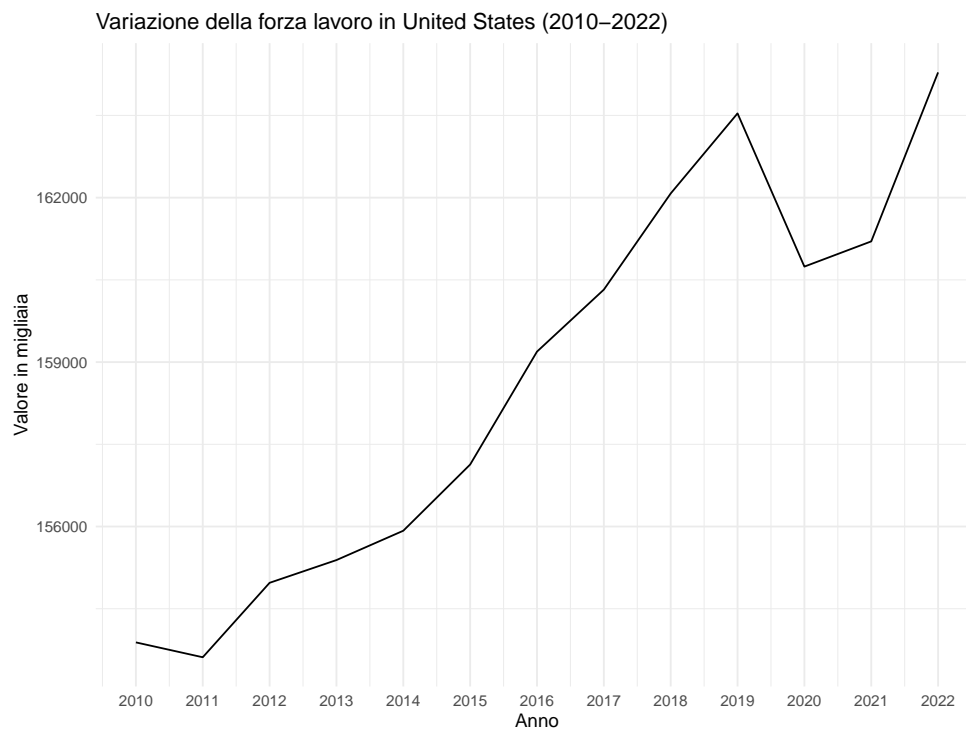
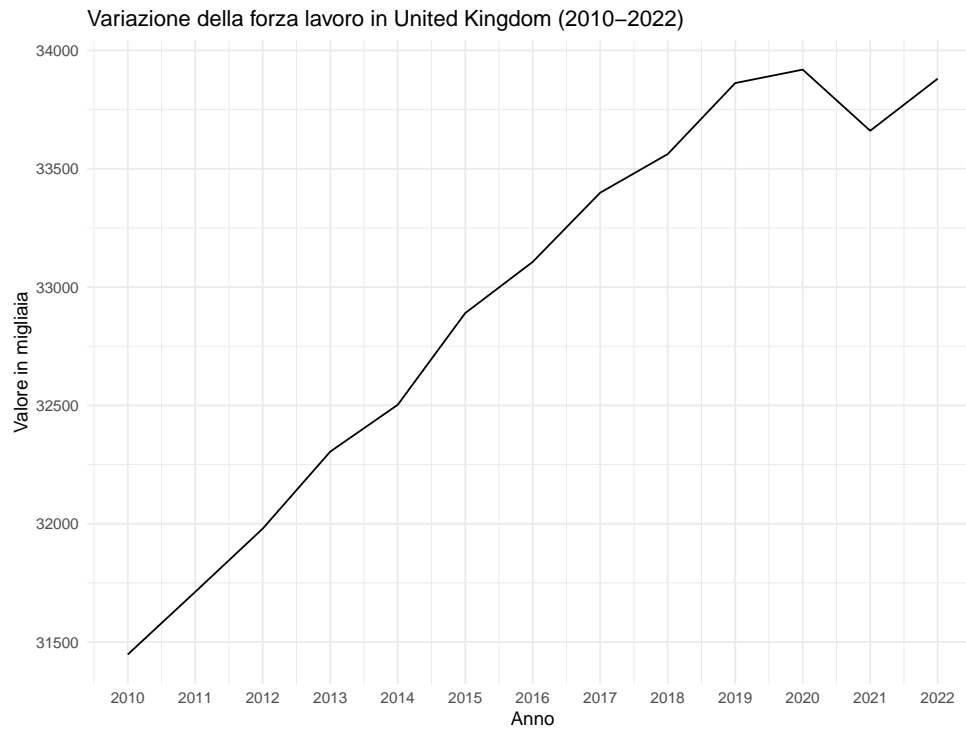


Variazione della forza lavoro in Switzerland (2010–2022)



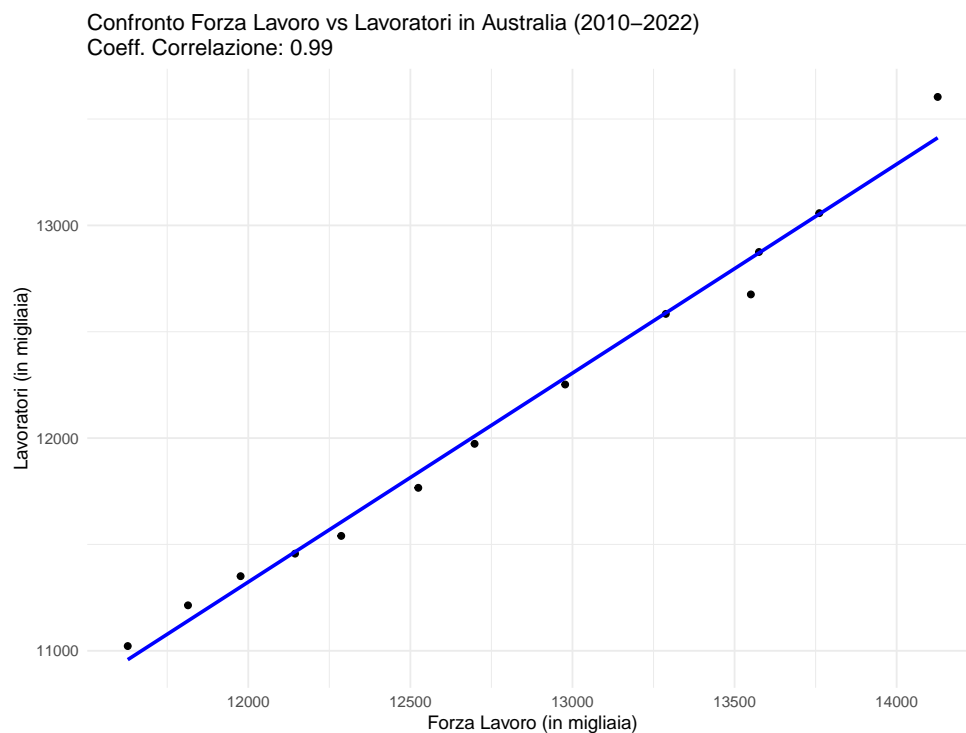
Variazione della forza lavoro in Türkiye (2010–2022)





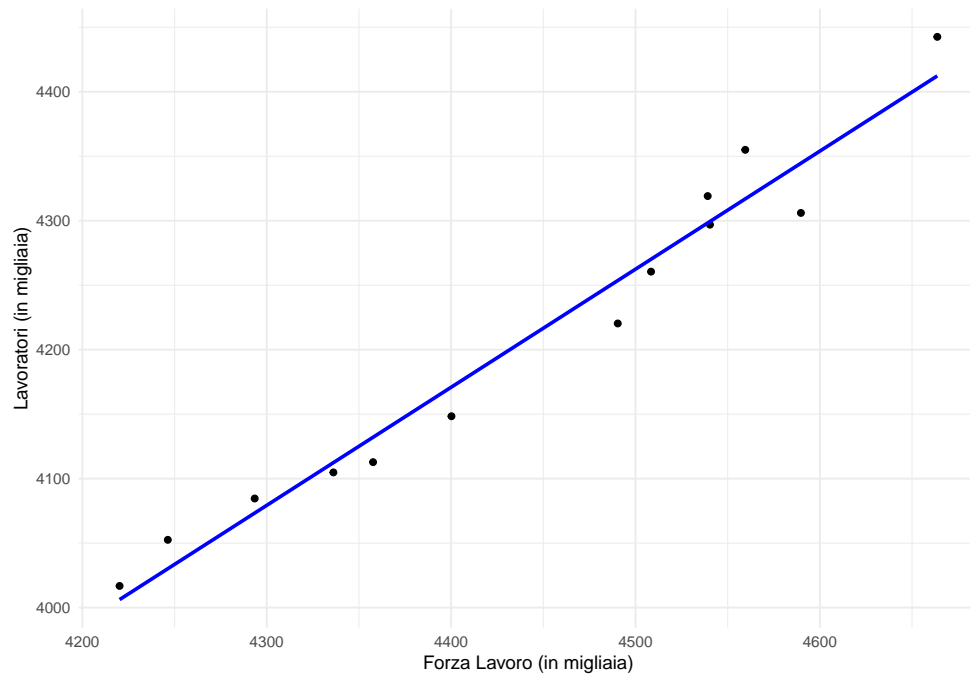
## 2.2 Scatterplot e Regressione Lineare

Per ogni paese, viene mostrato uno scatterplot che ha come variabile indipendente il numero di persone che costituiscono la forza lavoro (in migliaia), mentre come variabile dipendente si ha il numero di lavoratori (sempre in migliaia) dal 2010 al 2022. Questo tipo di analisi potrebbe aiutare a comprendere meglio la dinamica dell'occupazione e della forza lavoro in ogni paese nel corso del tempo.

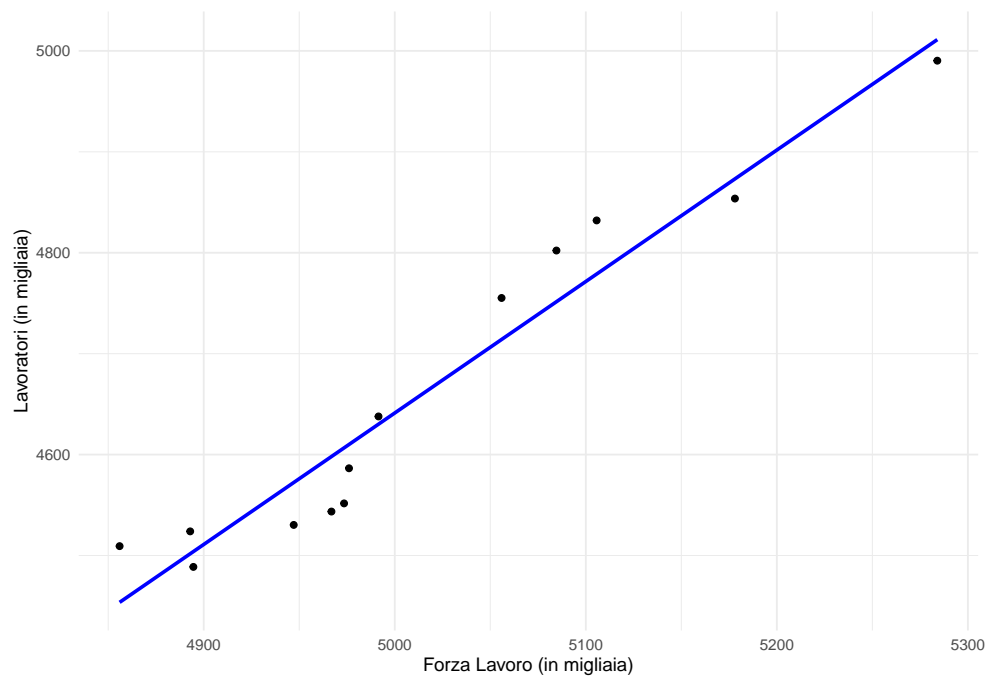




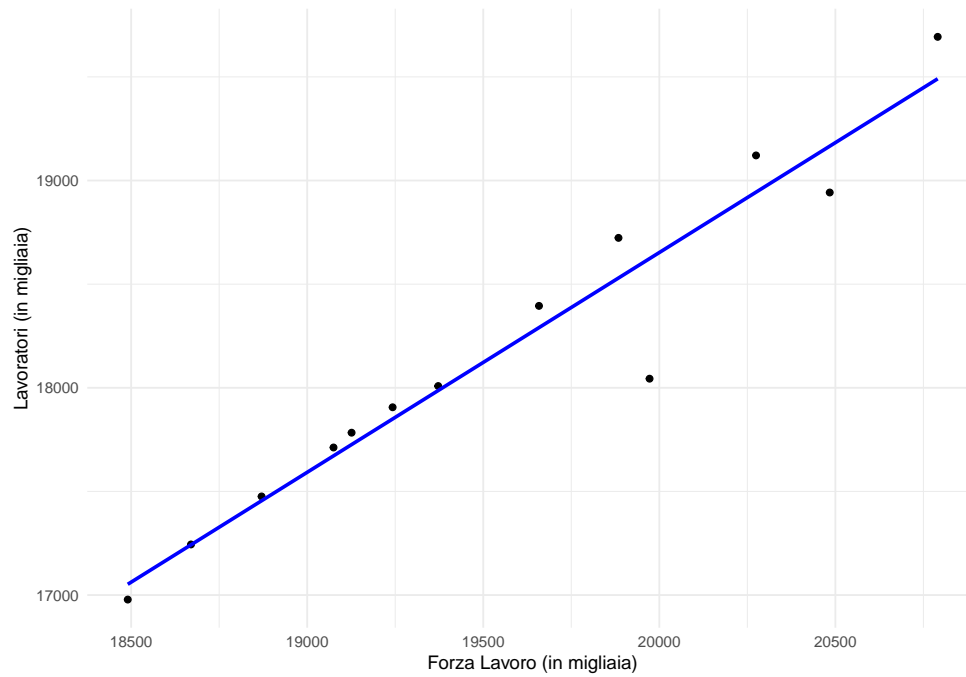
Confronto Forza Lavoro vs Lavoratori in Austria (2010–2022)  
Coeff. Correlazione: 0.98



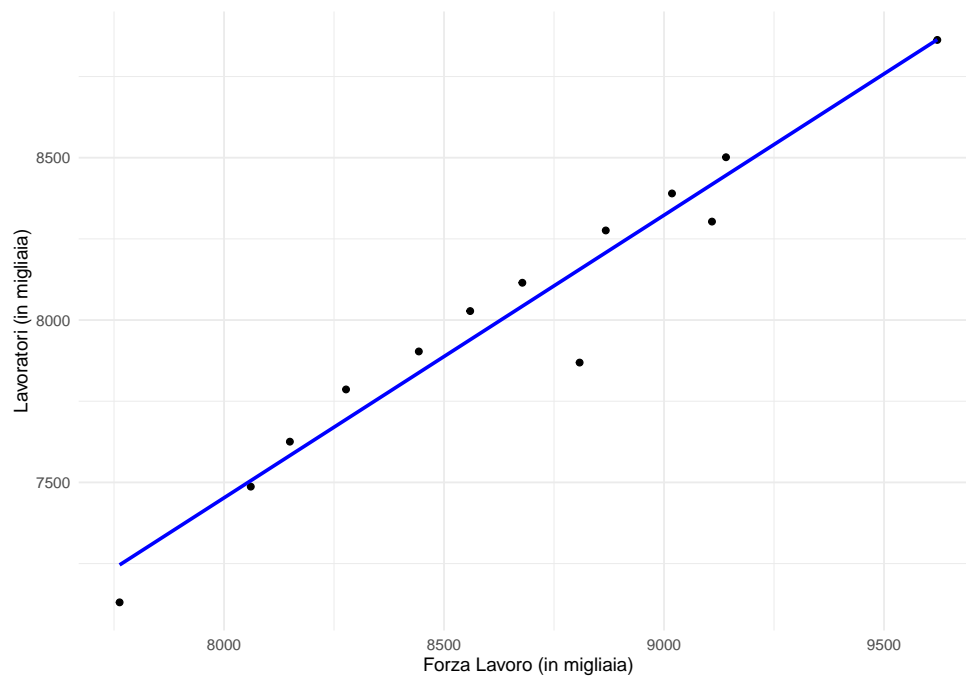
Confronto Forza Lavoro vs Lavoratori in Belgium (2010–2022)  
Coeff. Correlazione: 0.97

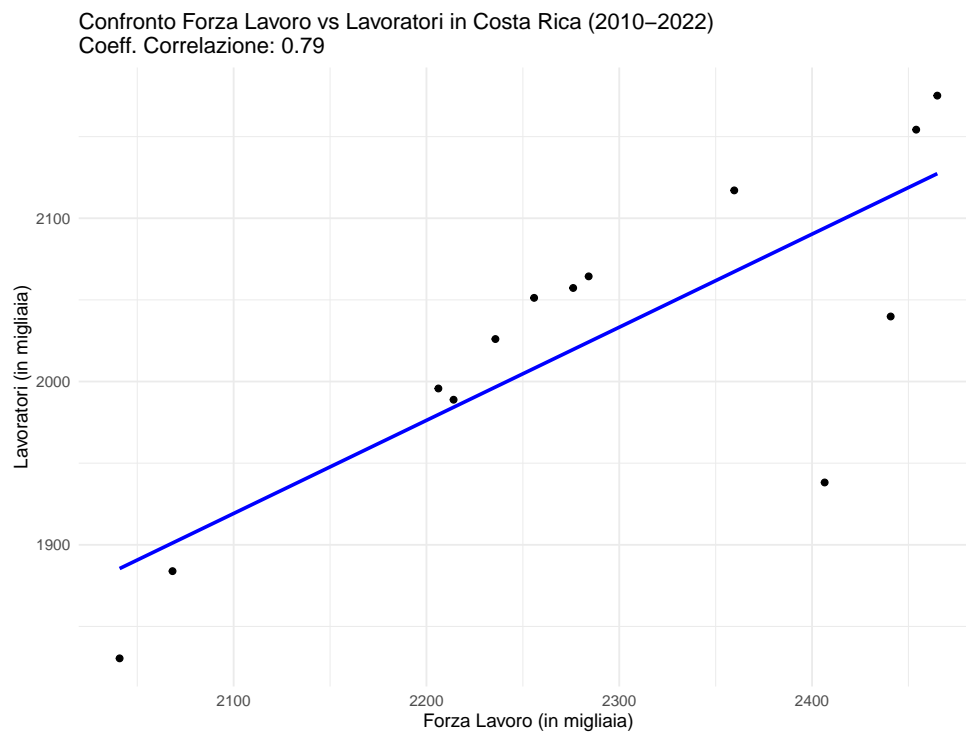
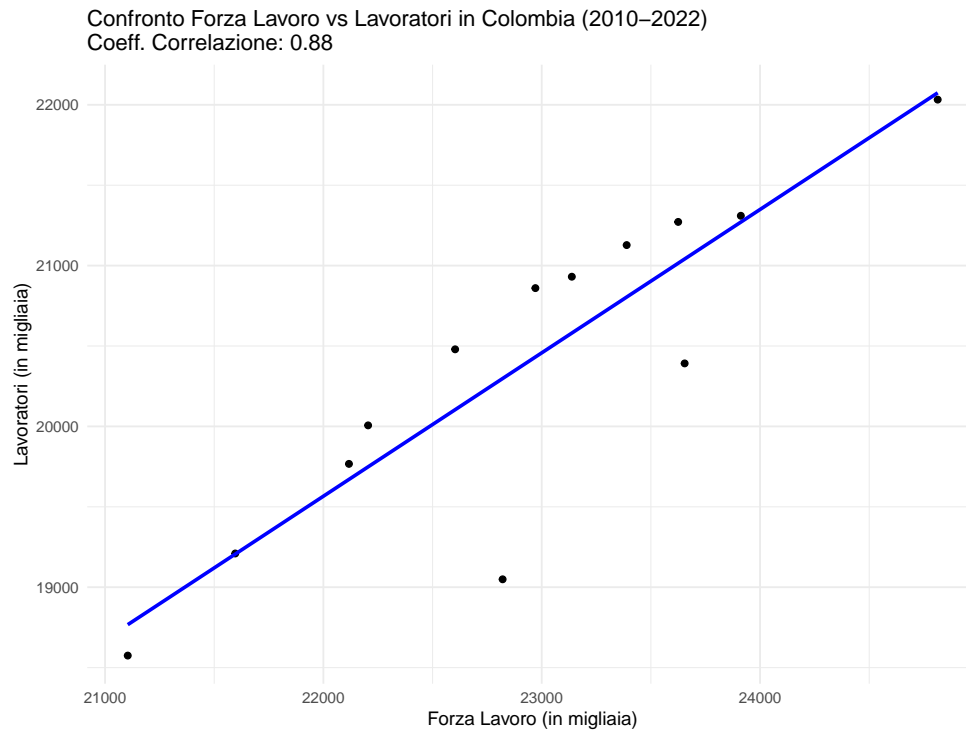


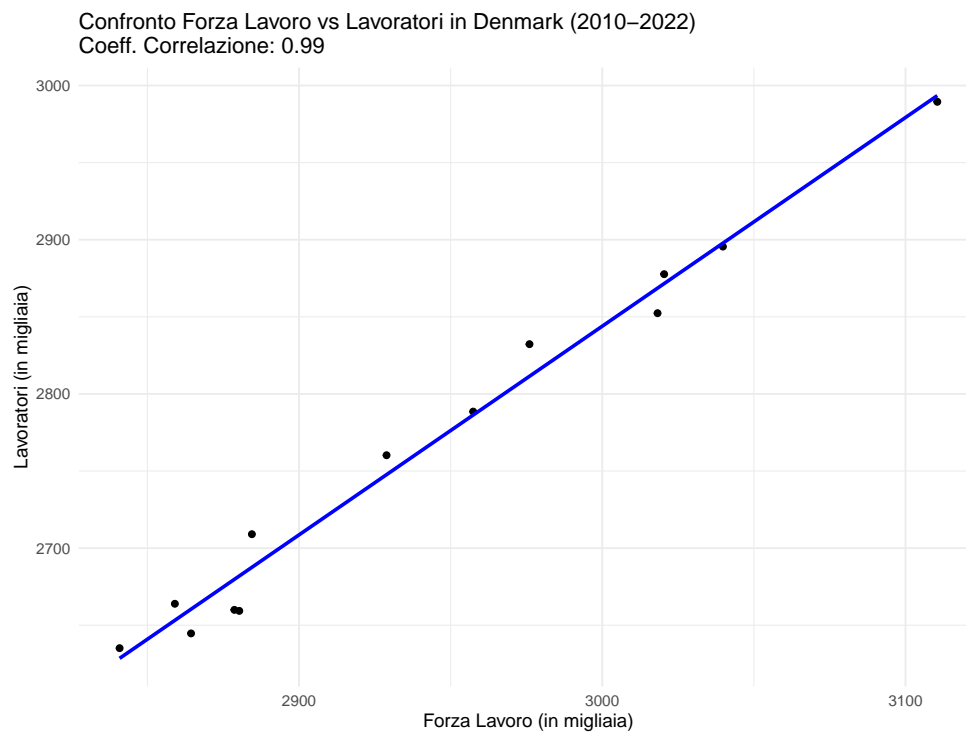
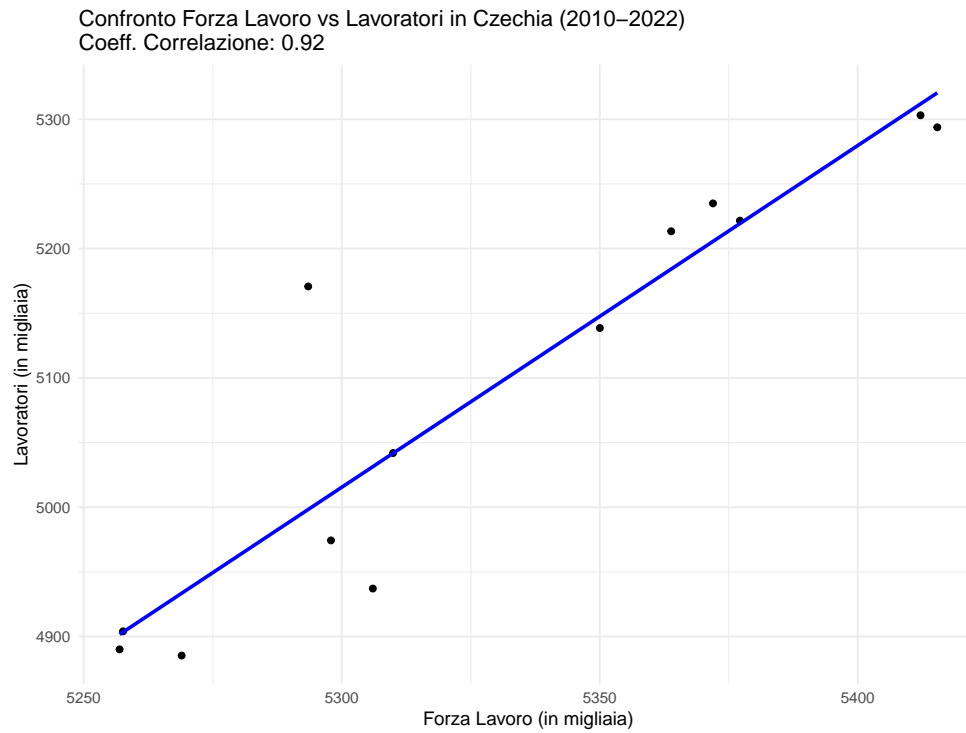
Confronto Forza Lavoro vs Lavoratori in Canada (2010–2022)  
Coeff. Correlazione: 0.96



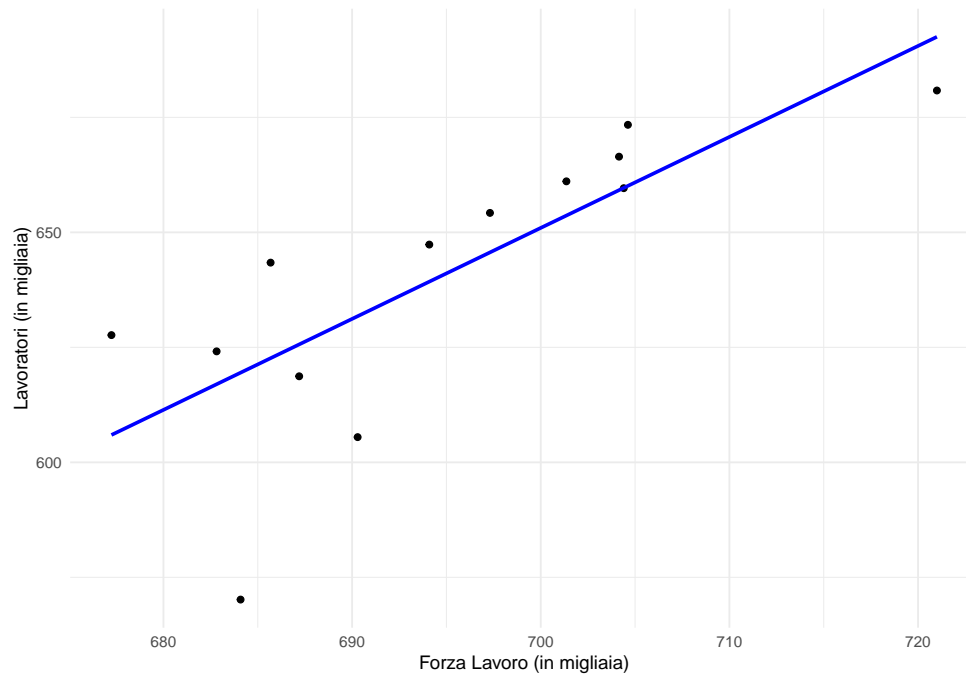
Confronto Forza Lavoro vs Lavoratori in Chile (2010–2022)  
Coeff. Correlazione: 0.97



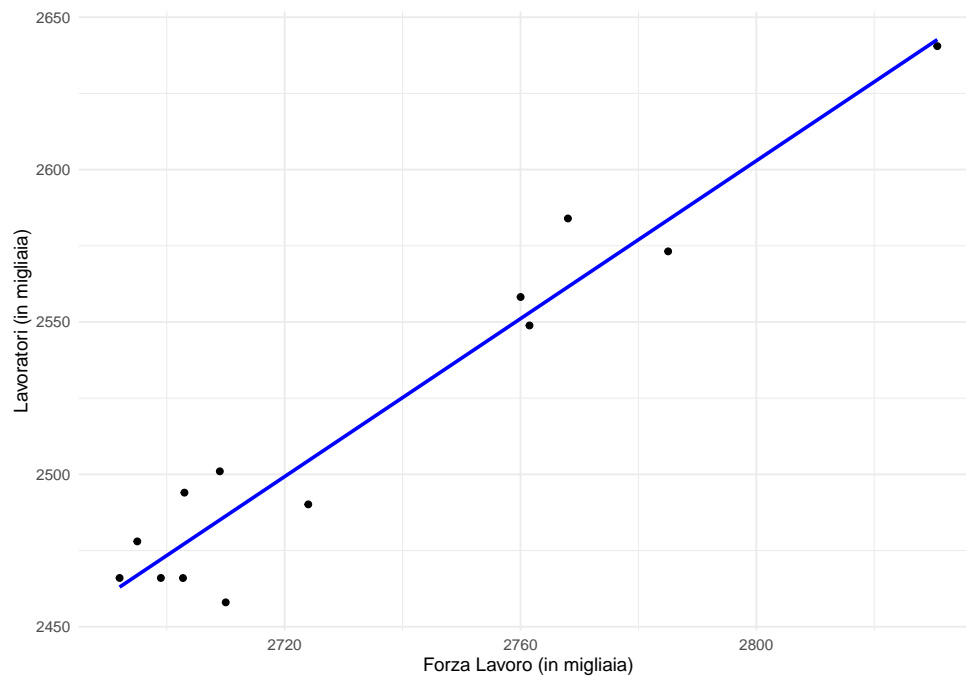




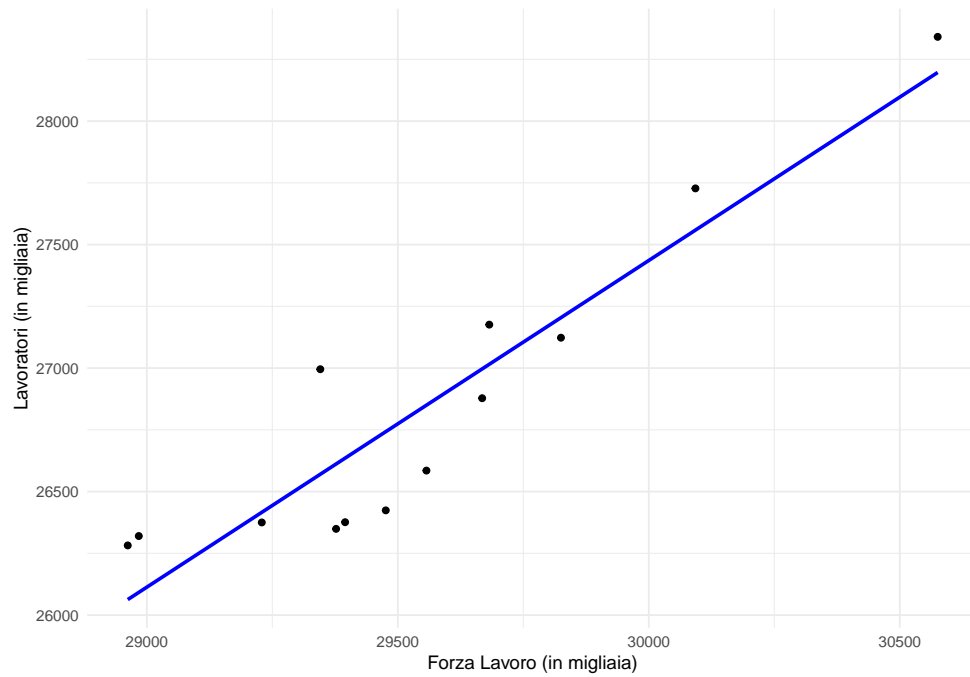
Confronto Forza Lavoro vs Lavoratori in Estonia (2010–2022)  
Coeff. Correlazione: 0.77



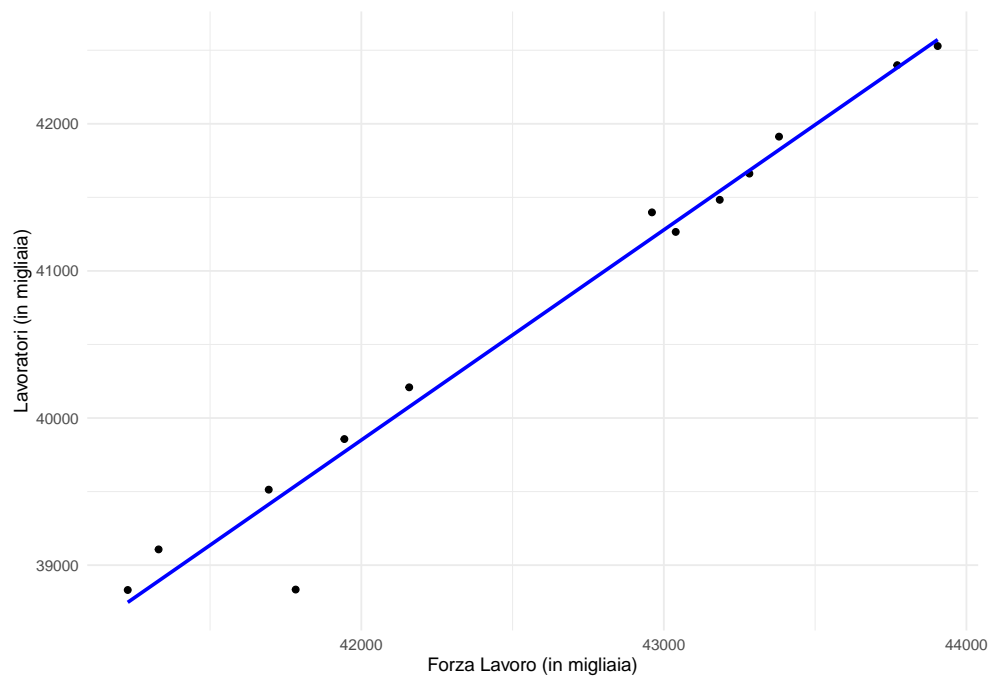
Confronto Forza Lavoro vs Lavoratori in Finland (2010–2022)  
Coeff. Correlazione: 0.97

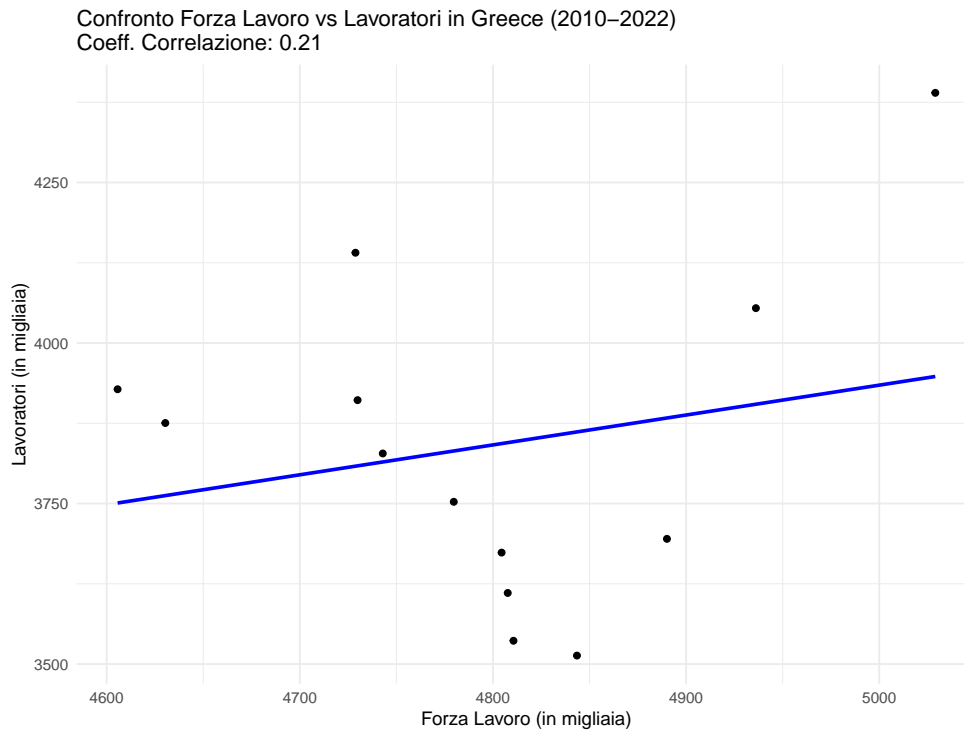


Confronto Forza Lavoro vs Lavoratori in France (2010–2022)  
Coeff. Correlazione: 0.92

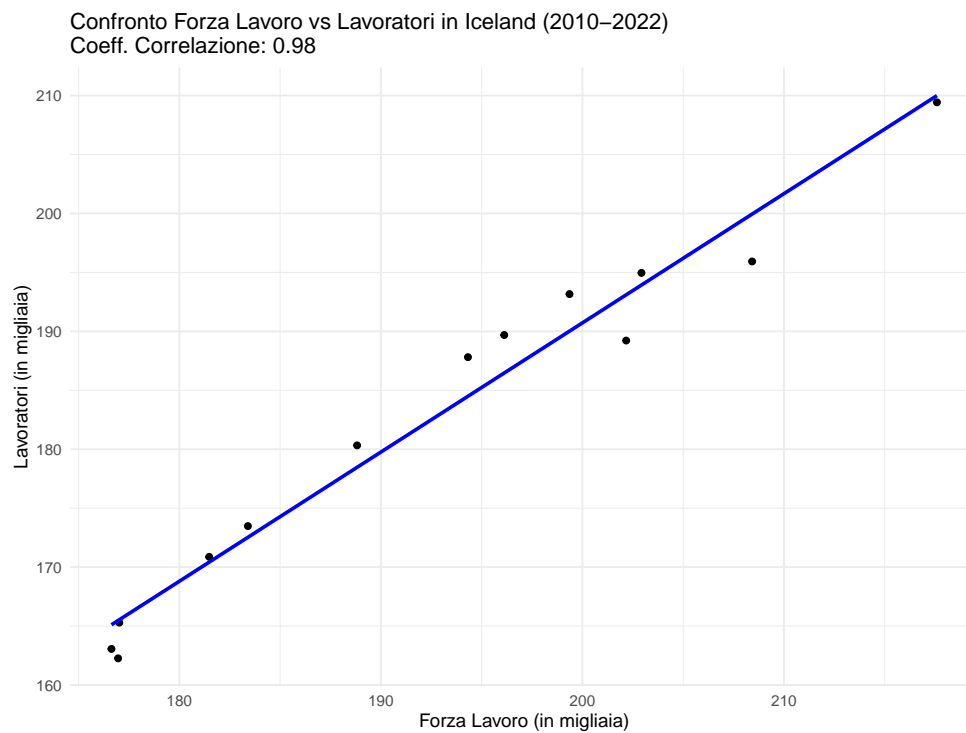
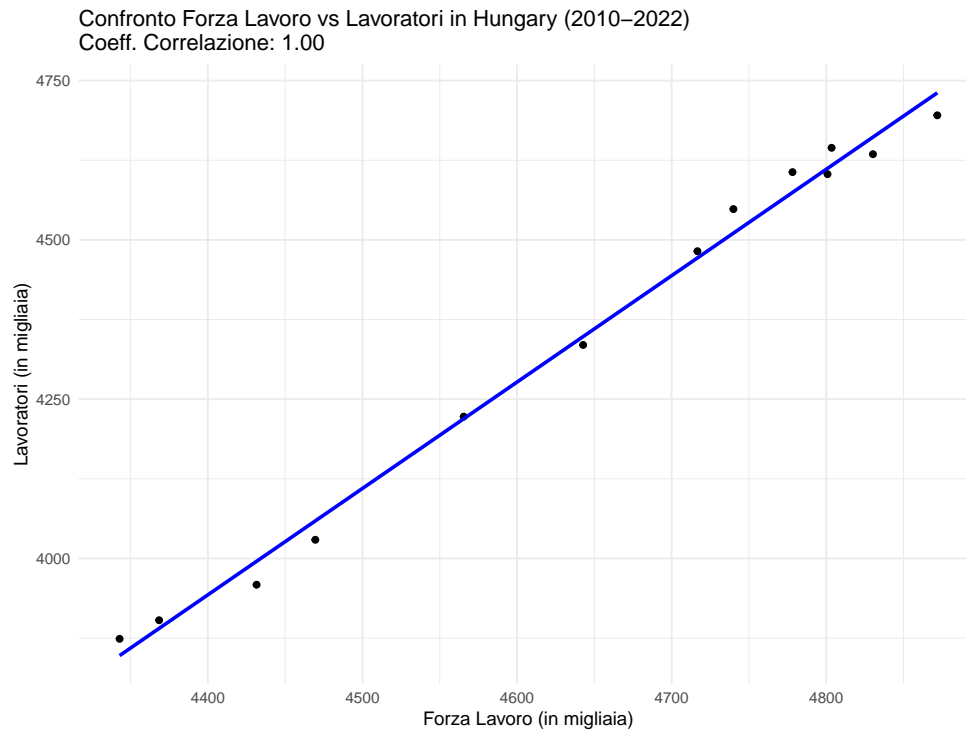


Confronto Forza Lavoro vs Lavoratori in Germany (2010–2022)  
Coeff. Correlazione: 0.99

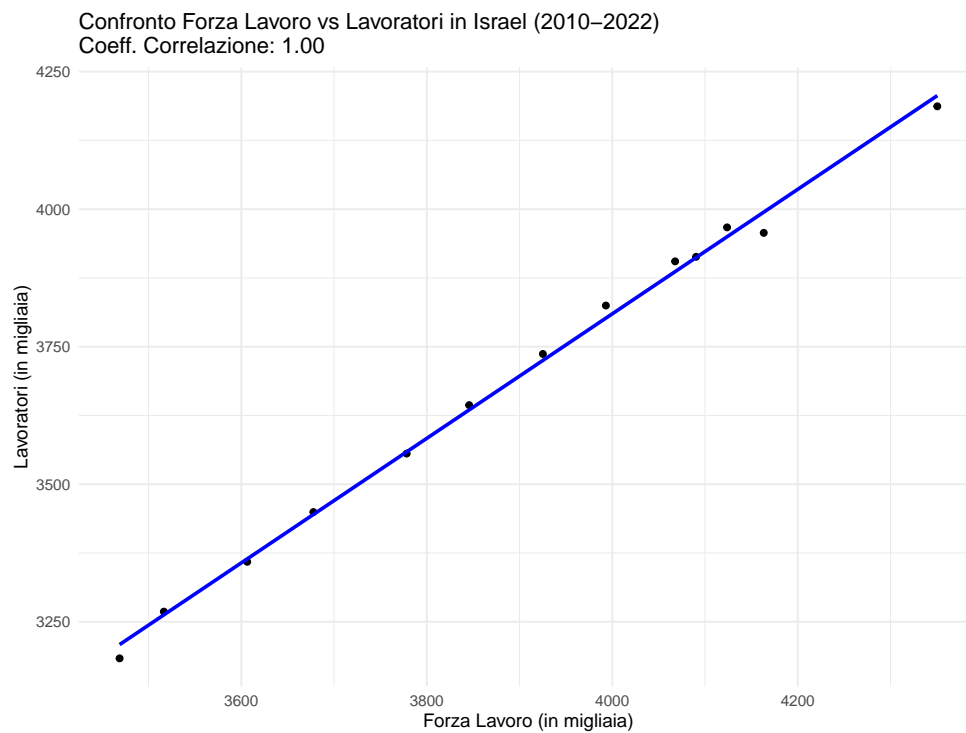
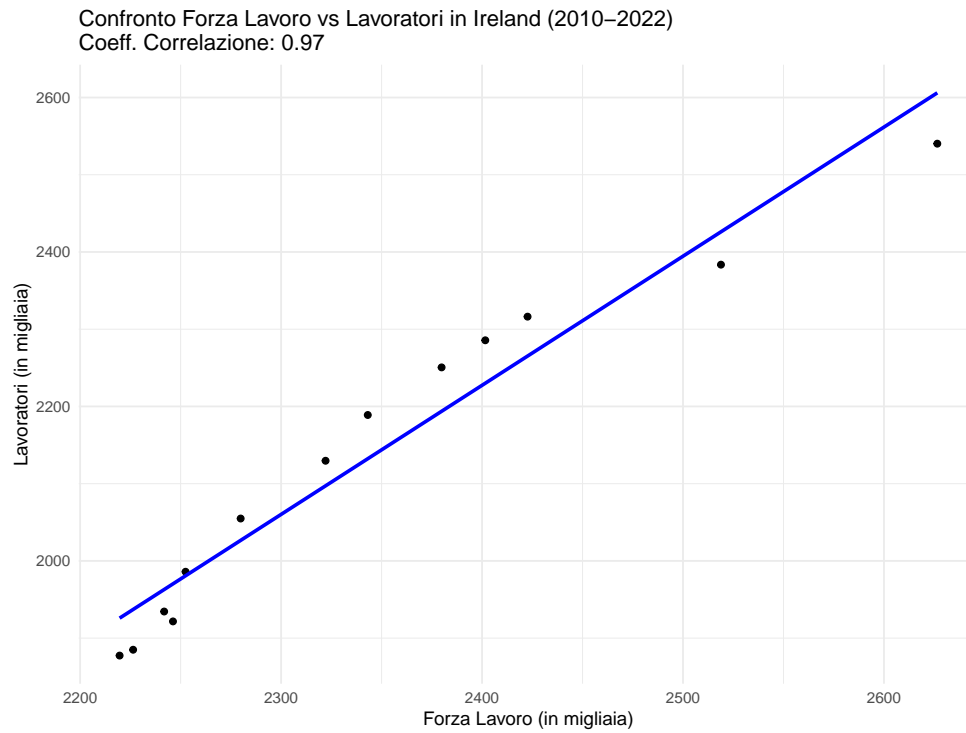


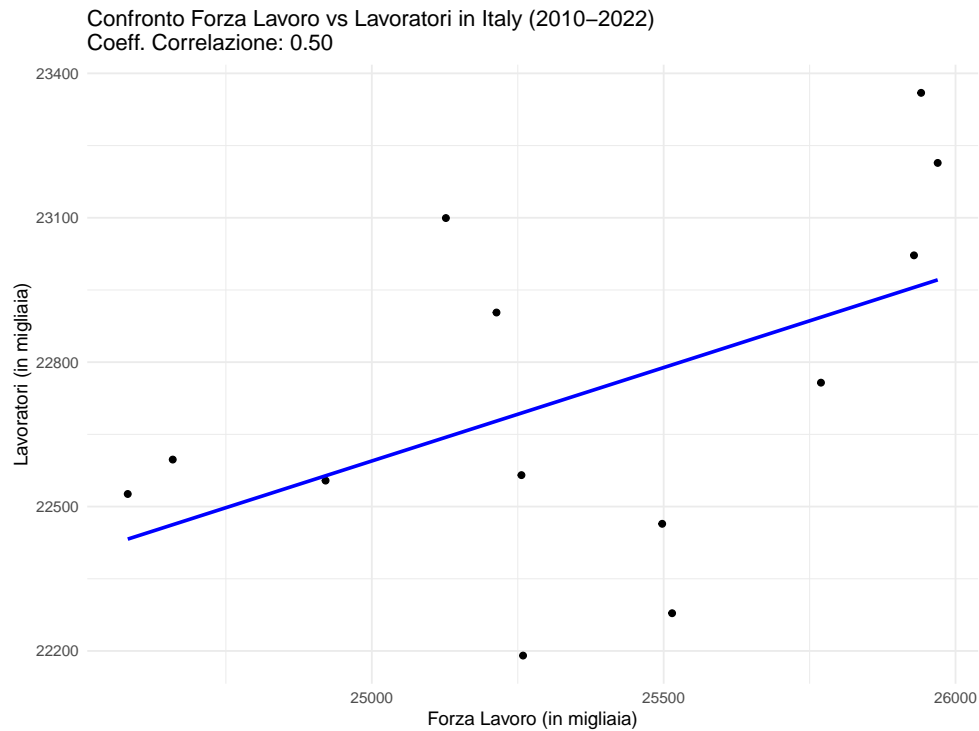


il coefficiente  $r$  di 0.21 per la Grecia suggerisce che c'è una leggera tendenza alla crescita nel numero di lavoratori all'aumentare del numero di persone nella forza lavoro. Tuttavia, la correlazione è considerata debole poiché il valore assoluto del coefficiente è abbastanza basso, il che indica che la relazione tra le due variabili non è particolarmente forte. Possiamo comunque ipotizzare che la crisi finanziaria del 2008 abbia avuto un impatto sull'occupazione in Grecia e potrebbe aver contribuito alla correlazione osservata. Tale crisi ha contribuito all'aumento della disoccupazione e alla riduzione delle opportunità lavorative. Per concludere con una nota positiva, in Grecia, nel corso degli anni, sta lentamente diminuendo il numero di disoccupati. [10]



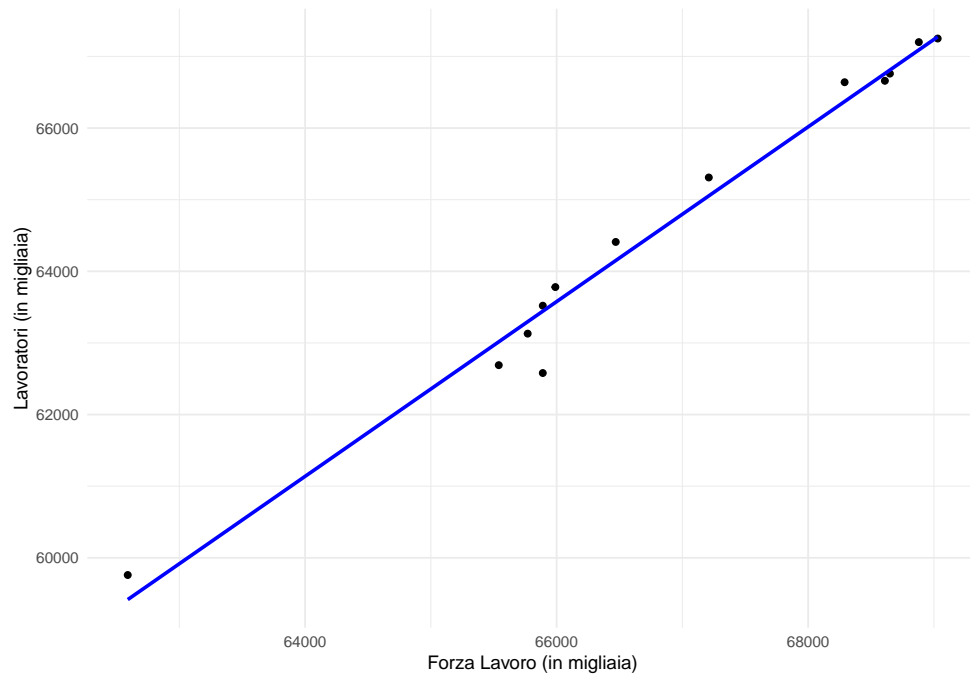




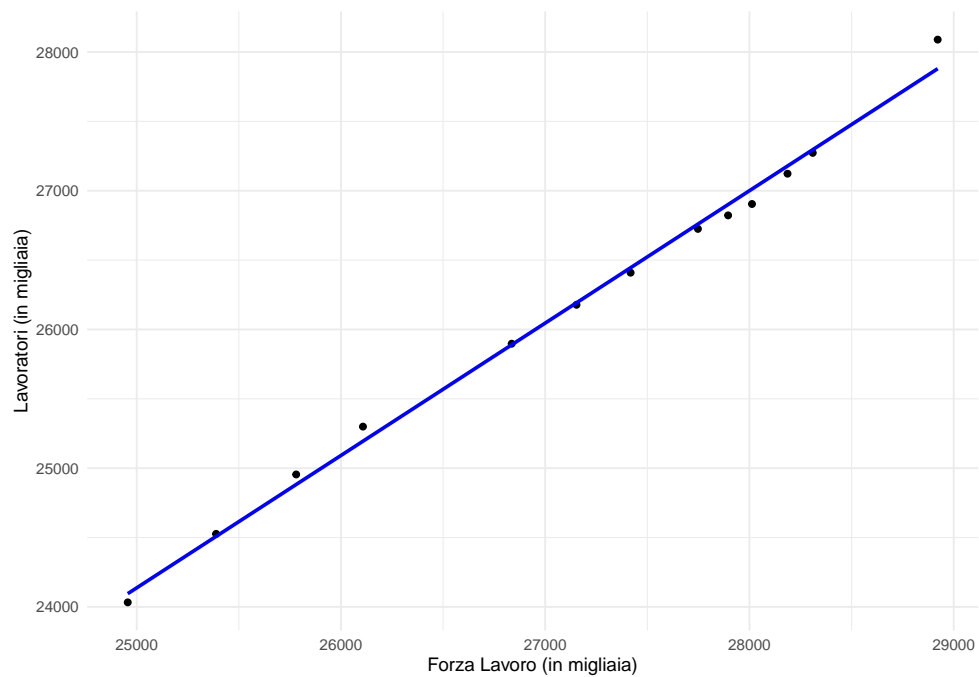


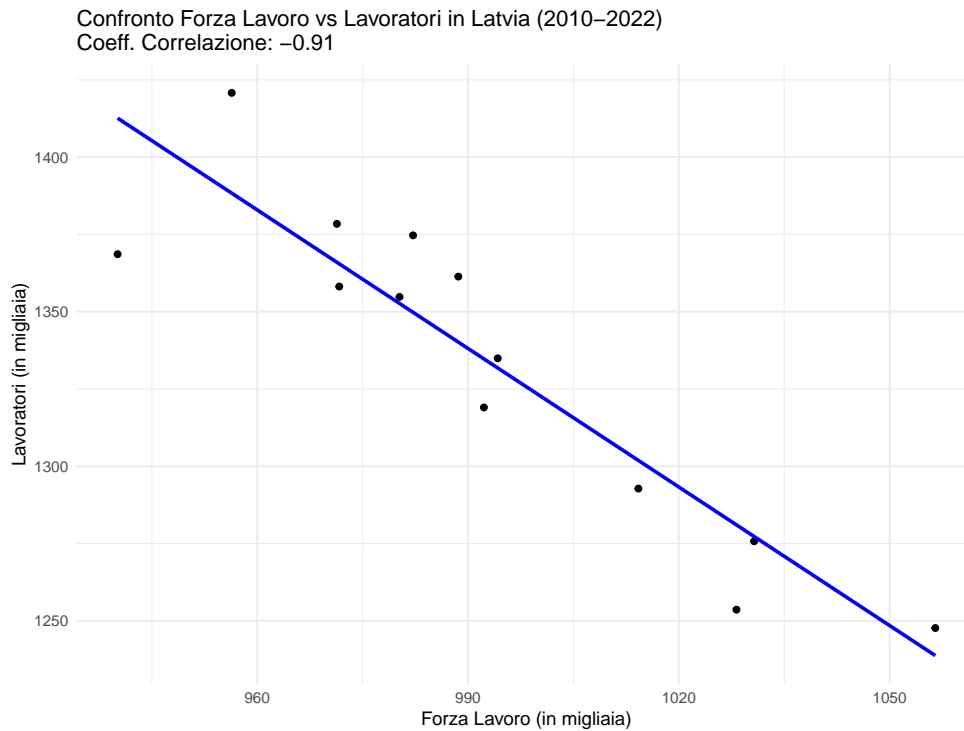
Un coefficiente di correlazione di 0.50 indica una correlazione positiva moderata tra le due variabili. In questo caso specifico, il coefficiente  $r$  di 0.50 per l'Italia suggerisce che c'è una tendenza alla crescita nel numero di lavoratori all'aumentare del numero di persone nella forza lavoro, e questa relazione è moderatamente forte, anche se minore della maggior parte degli altri paesi.

Confronto Forza Lavoro vs Lavoratori in Japan (2010–2022)  
Coeff. Correlazione: 0.99



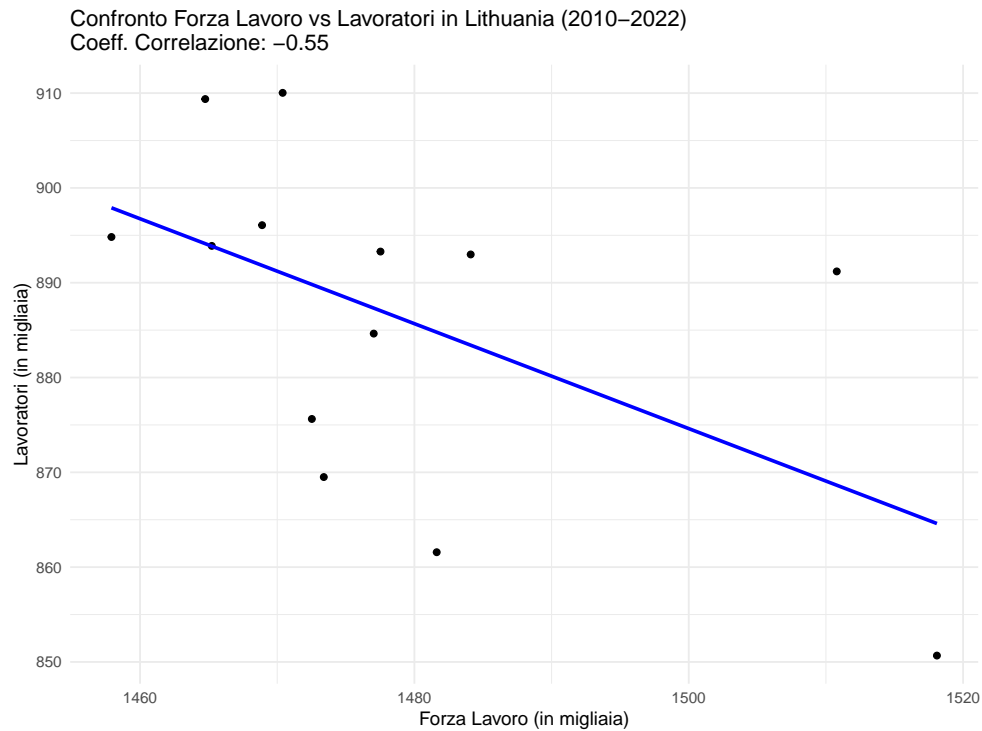
Confronto Forza Lavoro vs Lavoratori in Korea (2010–2022)  
Coeff. Correlazione: 1.00





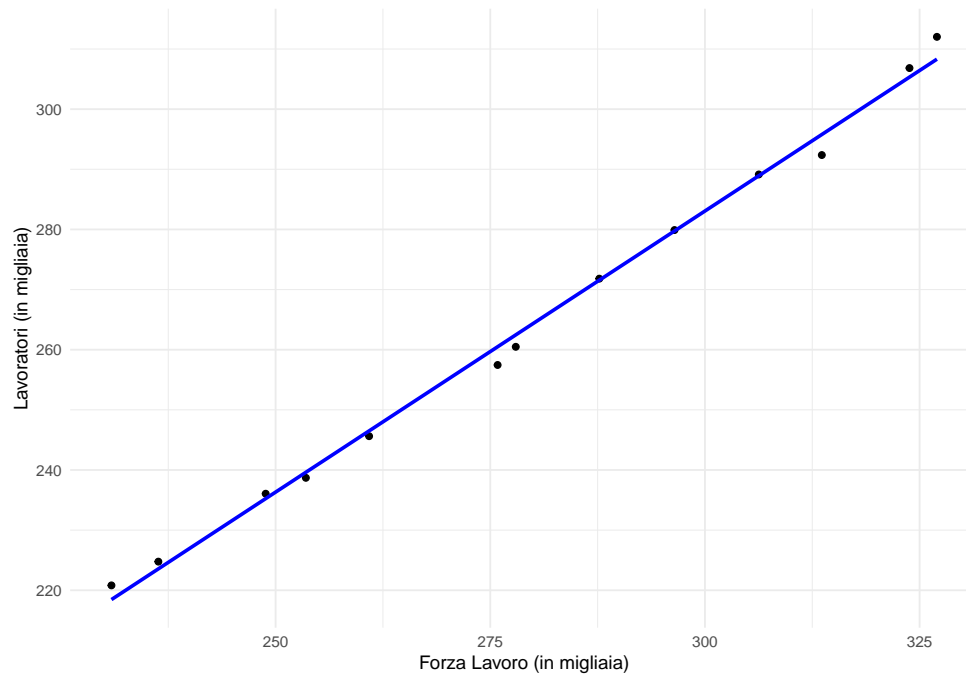
Un coefficiente di correlazione di  $-0.91$  (il più basso in assoluto fra tutti i paesi) per la Lettonia indica una forte correlazione negativa tra il numero di persone nella forza lavoro e il numero di lavoratori effettivamente impiegati nel periodo considerato. In altre parole, c'è una forte tendenza alla diminuzione nel numero di lavoratori all'aumentare del numero di persone nella forza lavoro. Quando il coefficiente di correlazione si avvicina a  $-1$ , indica una relazione lineare negativa molto forte tra le due variabili.

La Lettonia ha subito fortemente la crisi del 2008. Gli anni dal 2009 al 2011 sono stati per la Lettonia un periodo di pesantissime misure di austerità: tagli a stipendi pubblici e pensioni (quest'ultimo poi dichiarato incostituzionale), licenziamento di un terzo dei dipendenti pubblici e stipendi ridotti del 30% a quelli rimasti, drastica riduzione della spesa per welfare e sanità. [11]

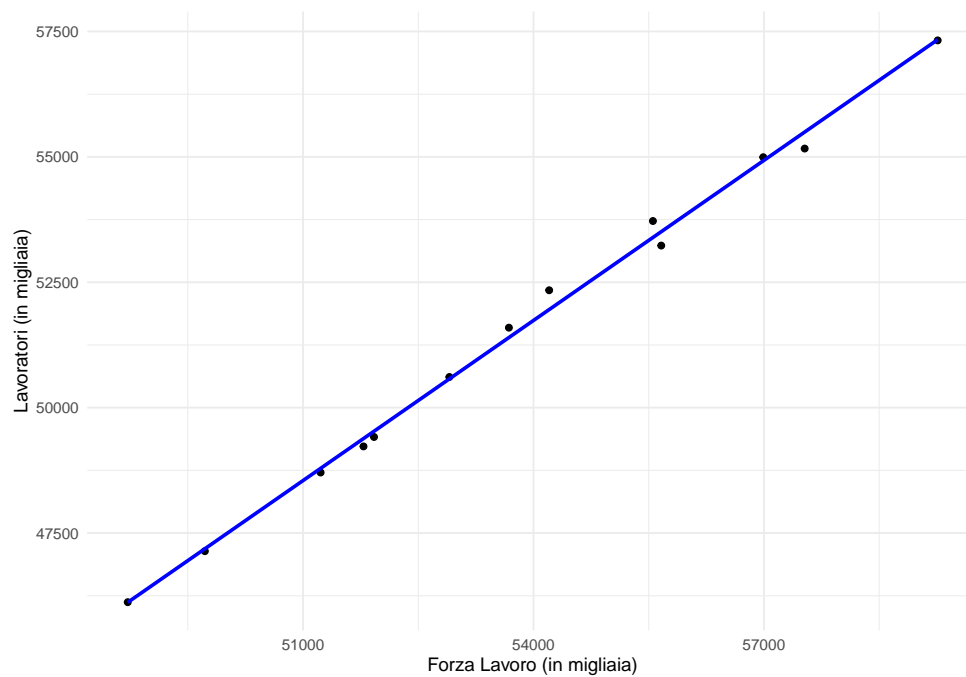


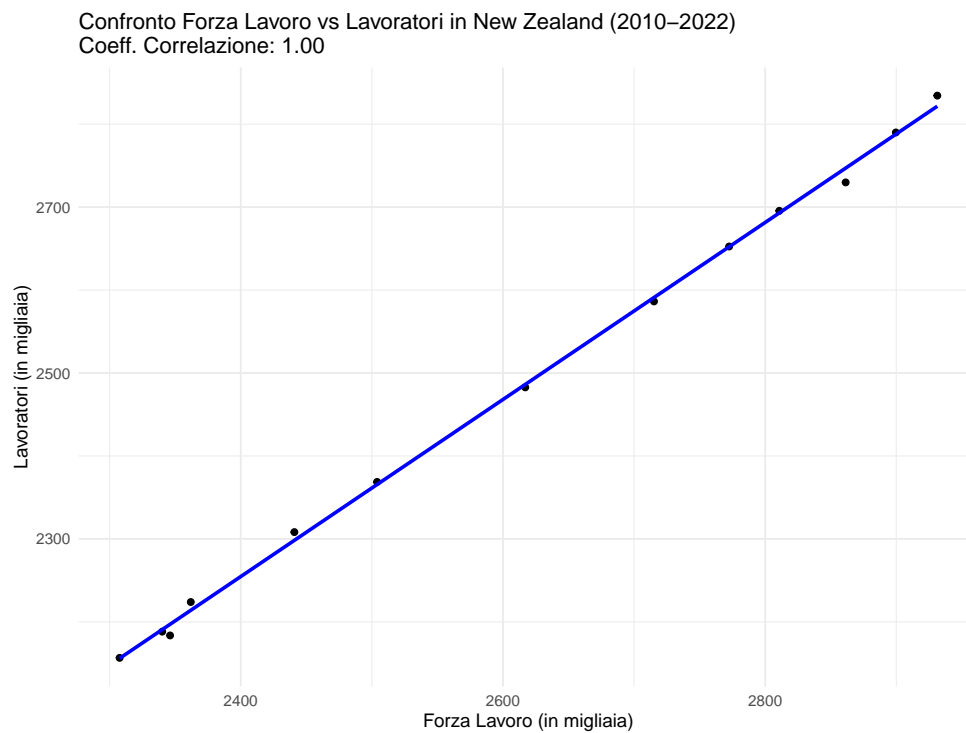
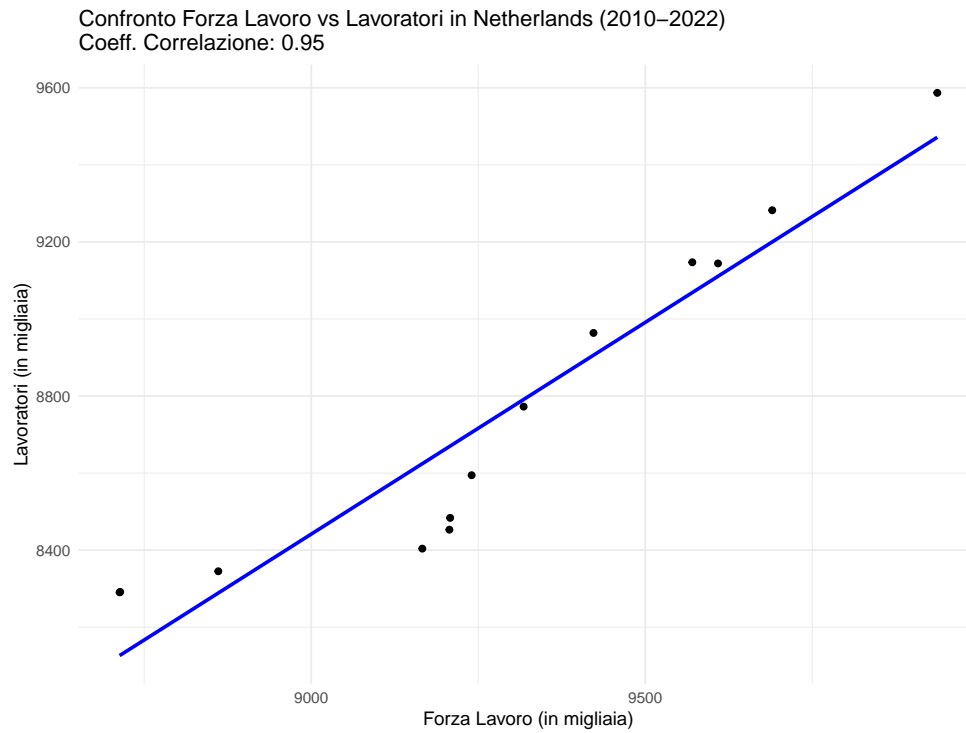
Un coefficiente di correlazione di  $-0.55$  per la Lituania indica una correlazione negativa moderata tra le due variabili. In questo contesto, il coefficiente di  $-0.55$  suggerisce che c'è una tendenza abbastanza significativa alla diminuzione nel numero di lavoratori all'aumentare del numero di persone nella forza lavoro. Poiché il coefficiente di  $-0.55$  è più vicino a  $-1$  rispetto al coefficiente della Grecia, indica una correlazione negativa più forte tra le due variabili in Lituania.

Confronto Forza Lavoro vs Lavoratori in Luxembourg (2010–2022)  
Coeff. Correlazione: 1.00

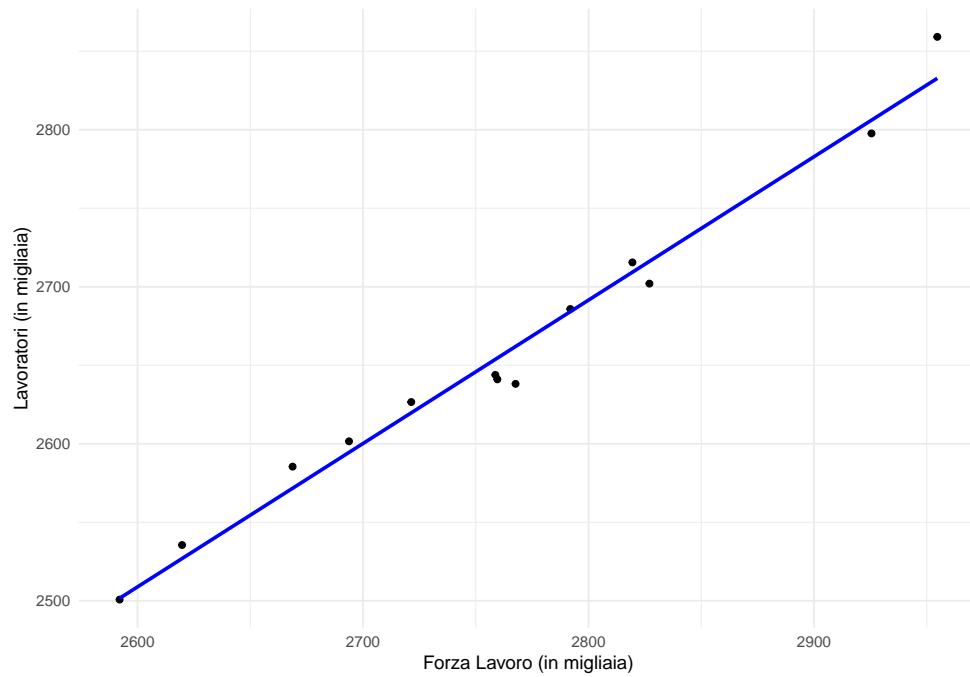


Confronto Forza Lavoro vs Lavoratori in Mexico (2010–2022)  
Coeff. Correlazione: 1.00

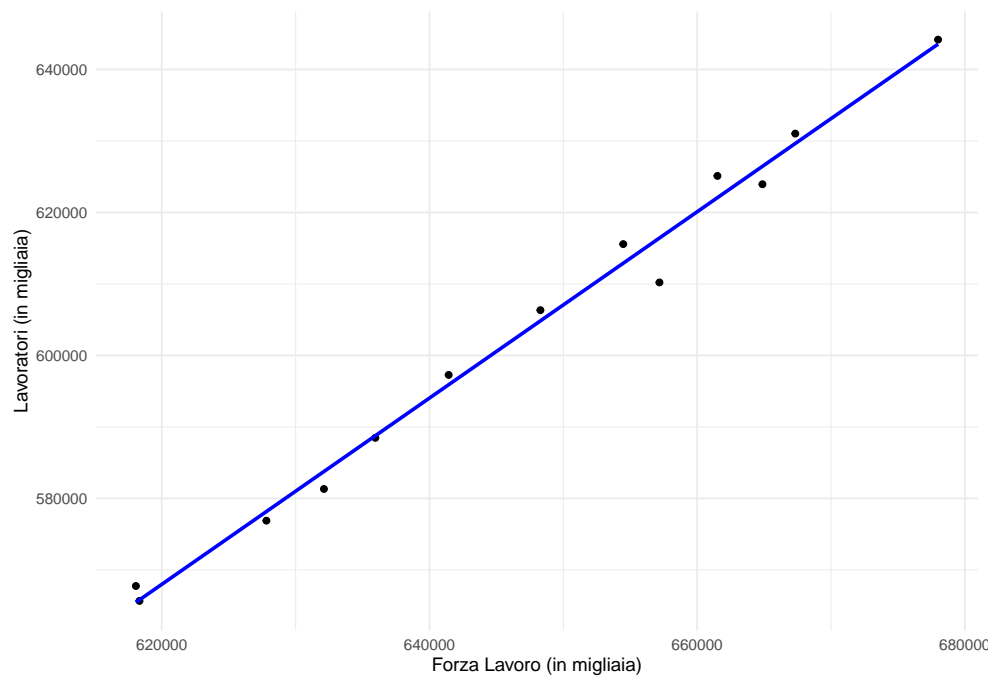




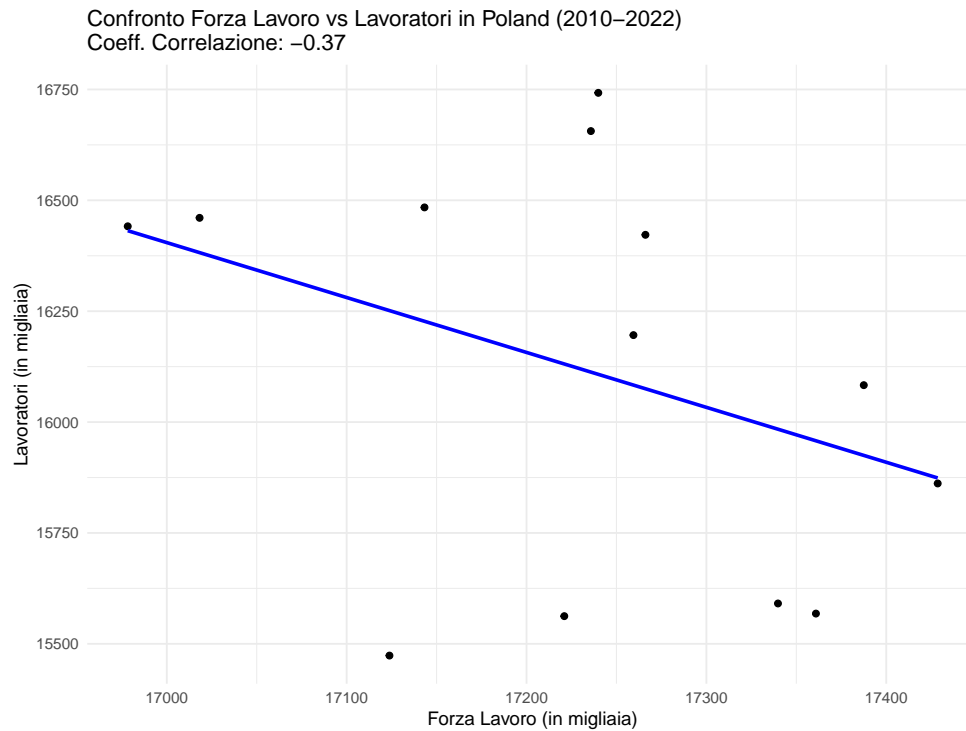
Confronto Forza Lavoro vs Lavoratori in Norway (2010–2022)  
Coeff. Correlazione: 0.99



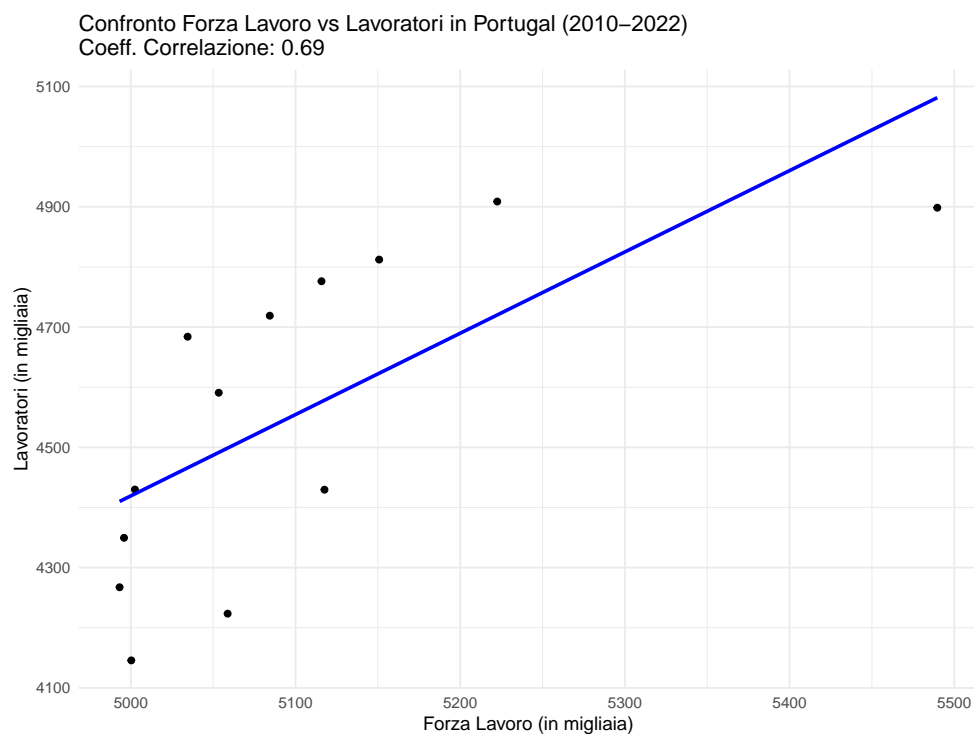
Confronto Forza Lavoro vs Lavoratori in OECD countries (2010–2022)  
Coeff. Correlazione: 0.99

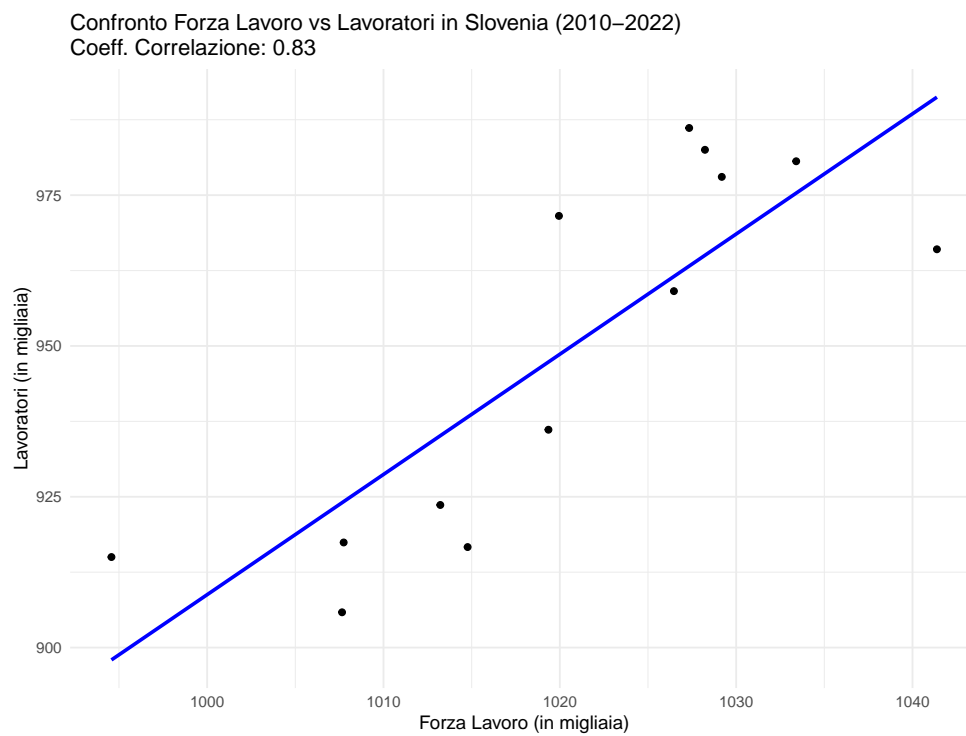
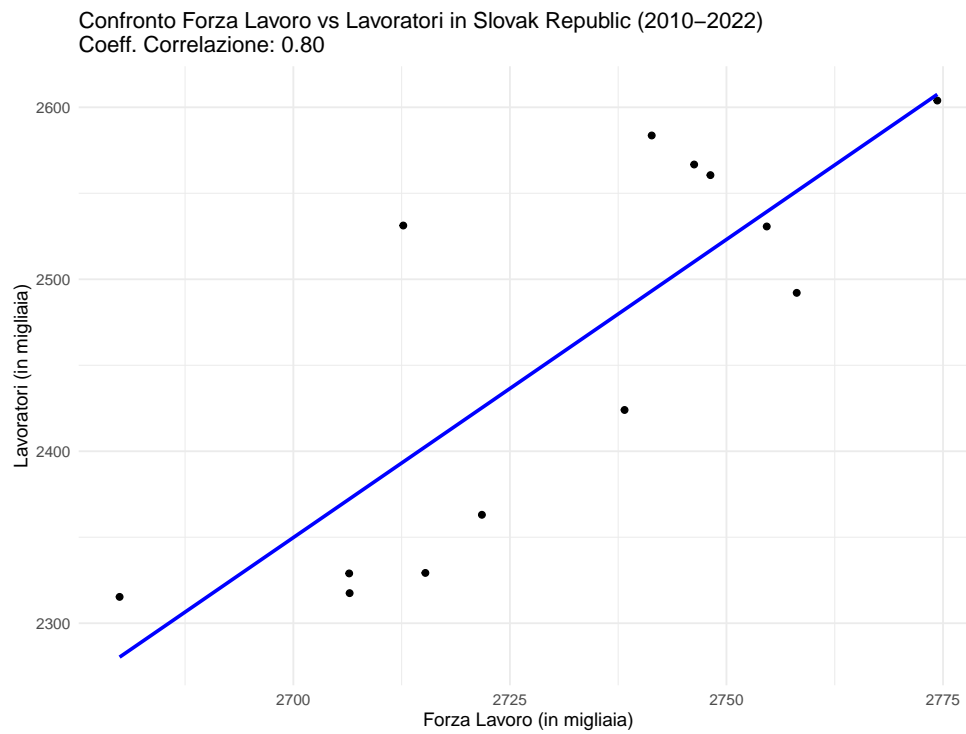


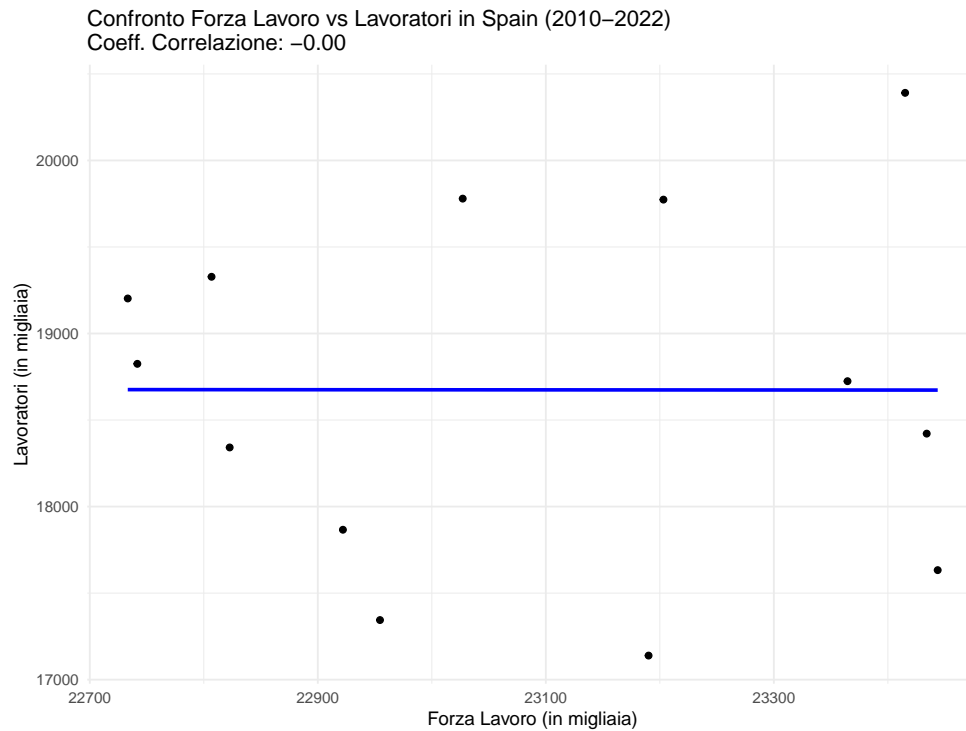




Un coefficiente di correlazione di  $-0.37$  per la Polonia indica una correlazione negativa moderata tra le due variabili. Sebbene il coefficiente sia più vicino a 0 rispetto a quello della Lituania, indica comunque una correlazione negativa tra le due variabili, anche se meno forte rispetto alla Lituania.

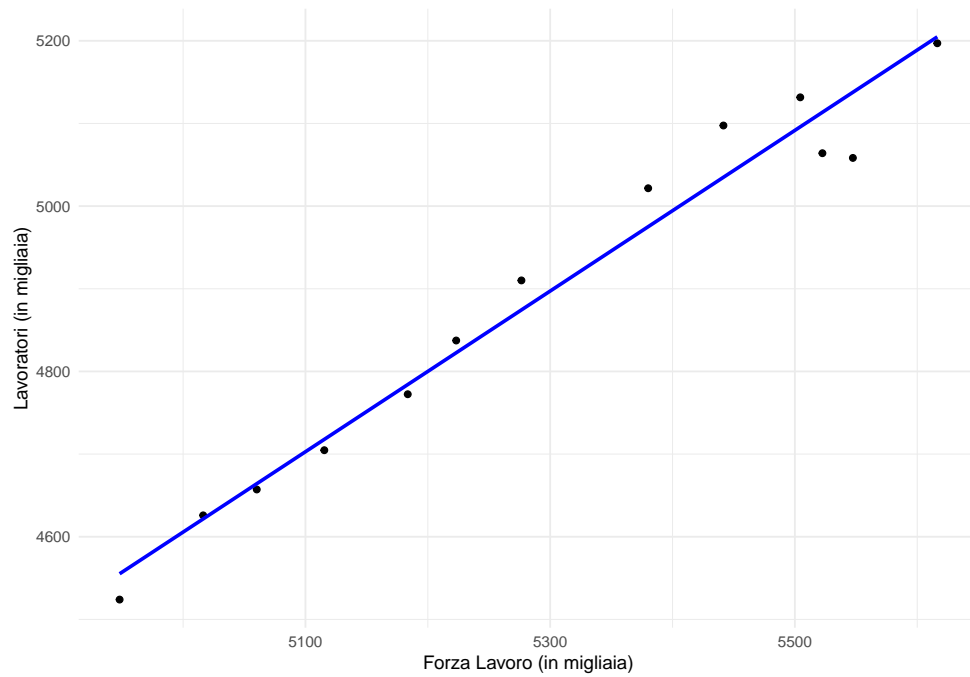




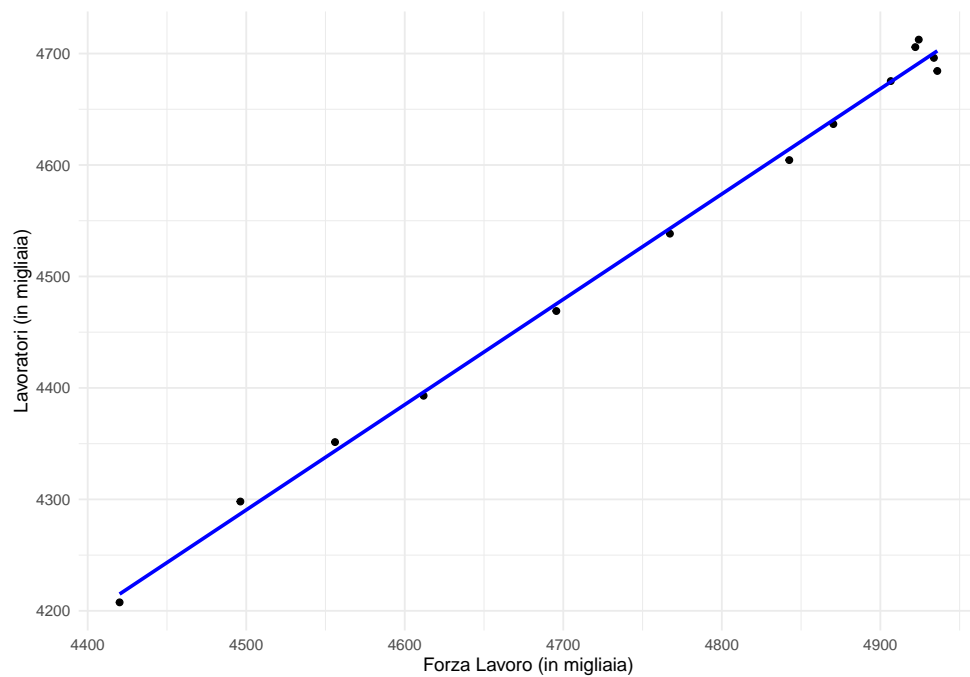


Un coefficiente di correlazione pari a 0 indica assenza di correlazione lineare tra le due variabili considerate. Nel caso della Spagna, un coefficiente di correlazione di 0 significa che non c'è una relazione lineare tra il numero di persone nella forza lavoro e il numero di lavoratori impiegati nel periodo considerato. La Grecia e la Spagna sono stati i paesi europei che più hanno risentito della crisi del 2008, infatti dal 2014 al 2022 sono stati i paesi con la percentuale di disoccupazione più alta in assoluto tra tutti e 38 i paesi che fanno parte degli OECD Countries. La presenza di una correlazione positiva debole (0.21) tra il numero di persone nella forza lavoro e il numero di lavoratori in Grecia potrebbe suggerire che la crisi economica abbia influenzato meno la Grecia rispetto alla Spagna in termini di occupazione.

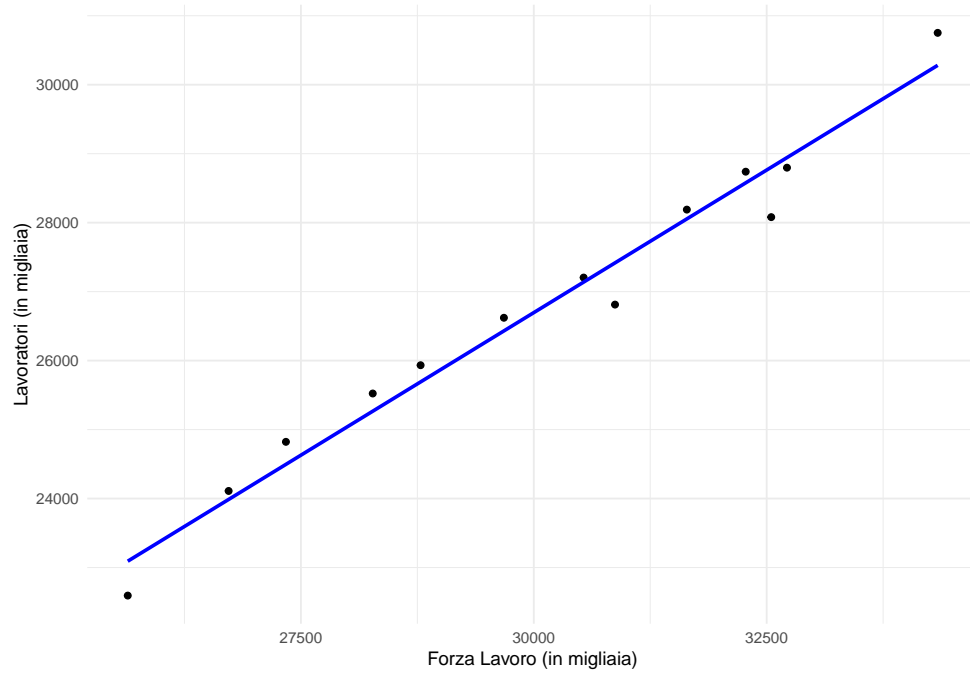
Confronto Forza Lavoro vs Lavoratori in Sweden (2010–2022)  
Coeff. Correlazione: 0.98



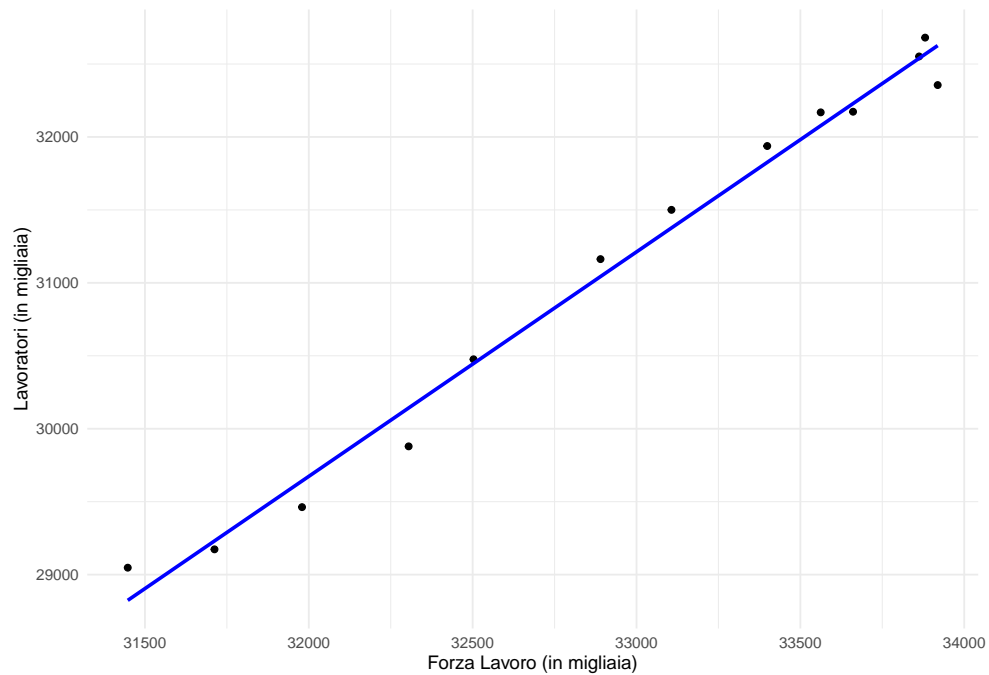
Confronto Forza Lavoro vs Lavoratori in Switzerland (2010–2022)  
Coeff. Correlazione: 1.00

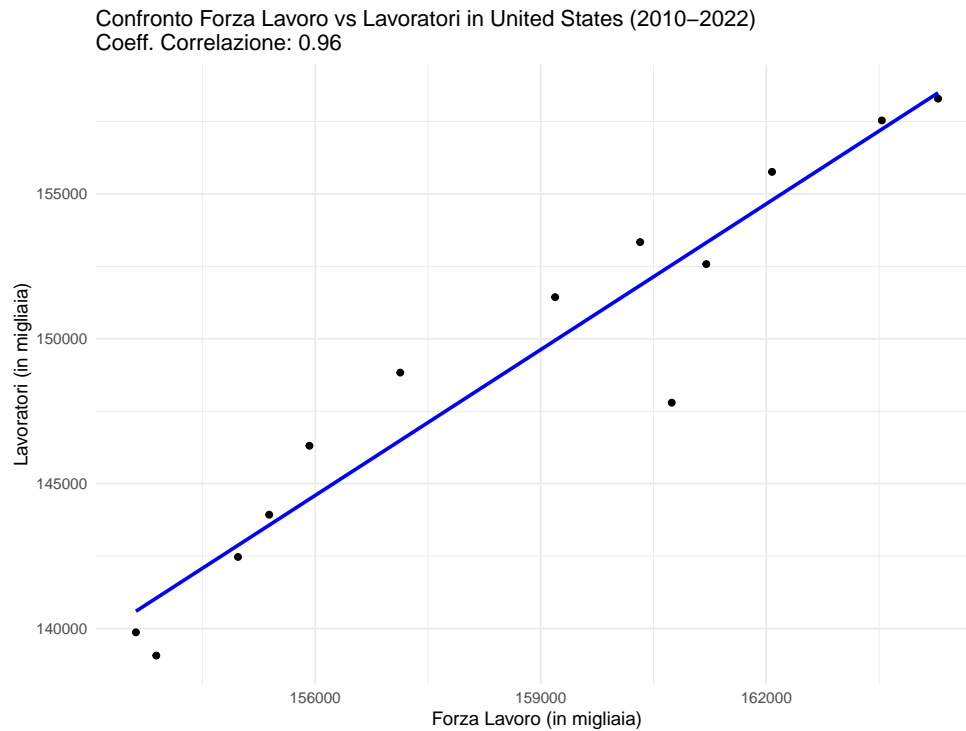


Confronto Forza Lavoro vs Lavoratori in Türkiye (2010–2022)  
Coeff. Correlazione: 0.99



Confronto Forza Lavoro vs Lavoratori in United Kingdom (2010–2022)  
Coeff. Correlazione: 0.99





Ricapitolando, I paesi che hanno un coefficiente  $r$  compreso tra -1 e 0 sono: Lettonia, Lituania, Spagna e Polonia. Questi paesi sono stati duramente colpiti dalla crisi del 2008, anche se la Polonia ha registrato un tasso di crescita economica più resiliente grazie a politiche economiche prudenti e a una maggiore stabilità finanziaria.

## CAPITOLO 3

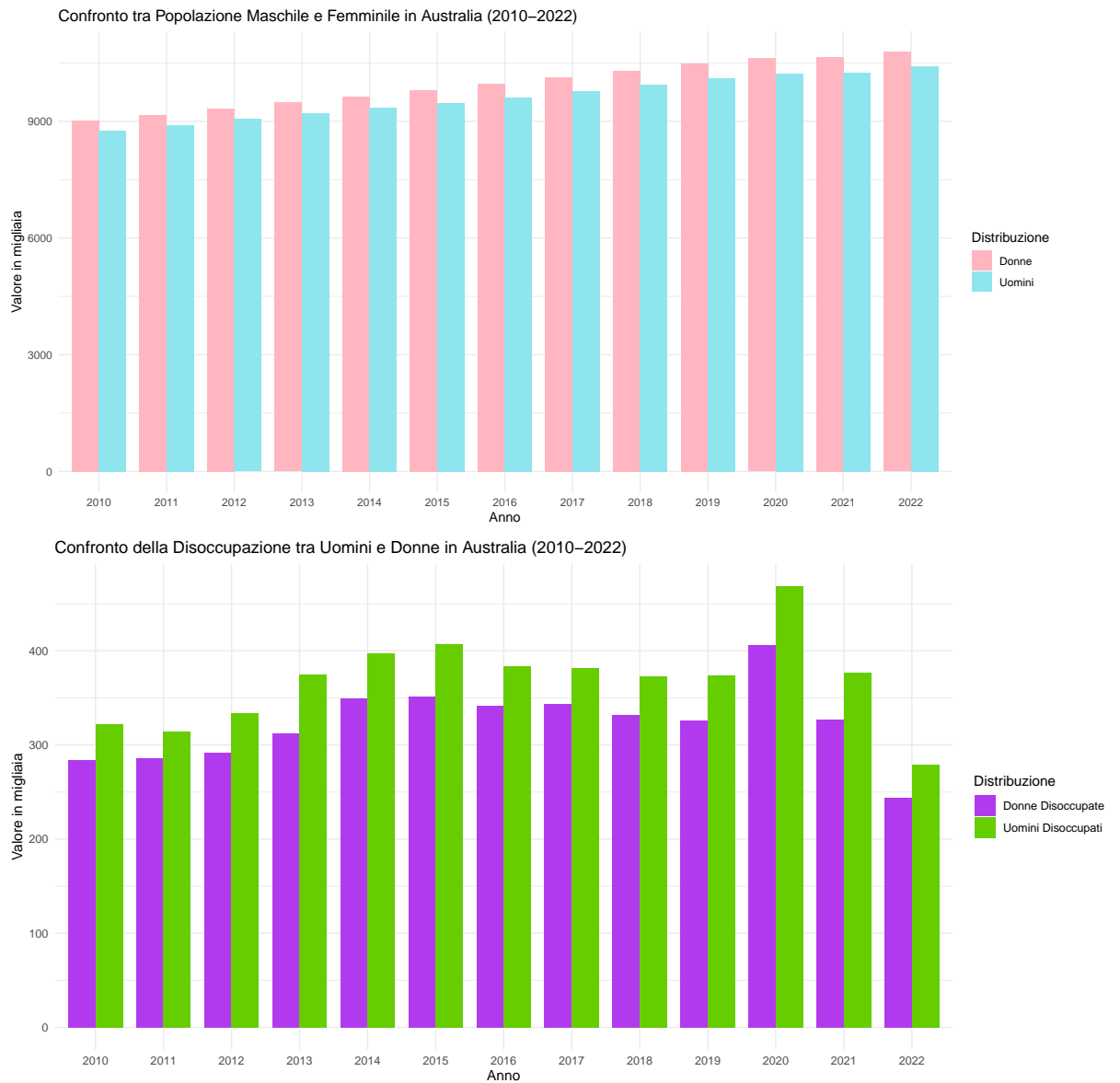
---

### Barplot

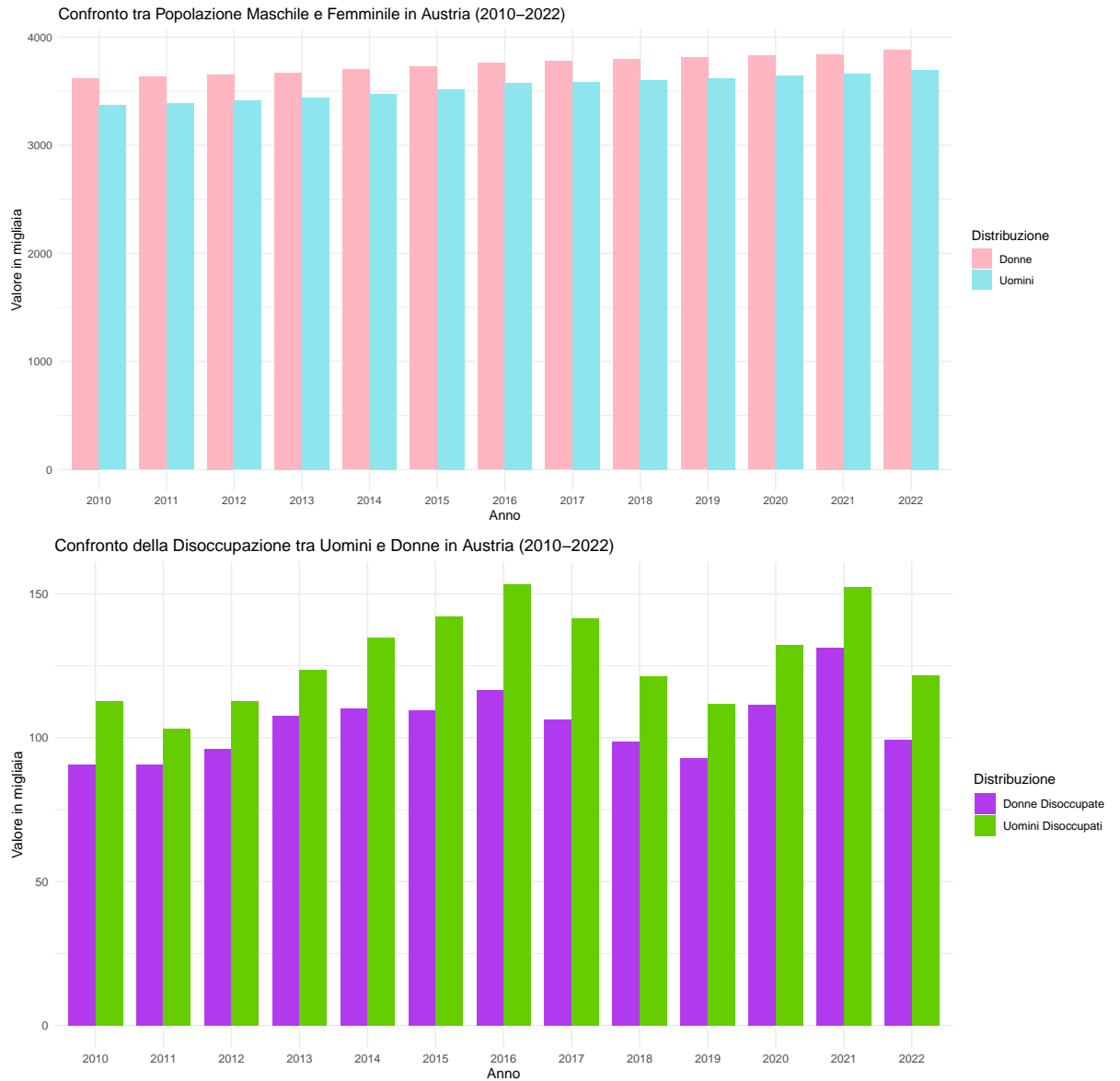
---

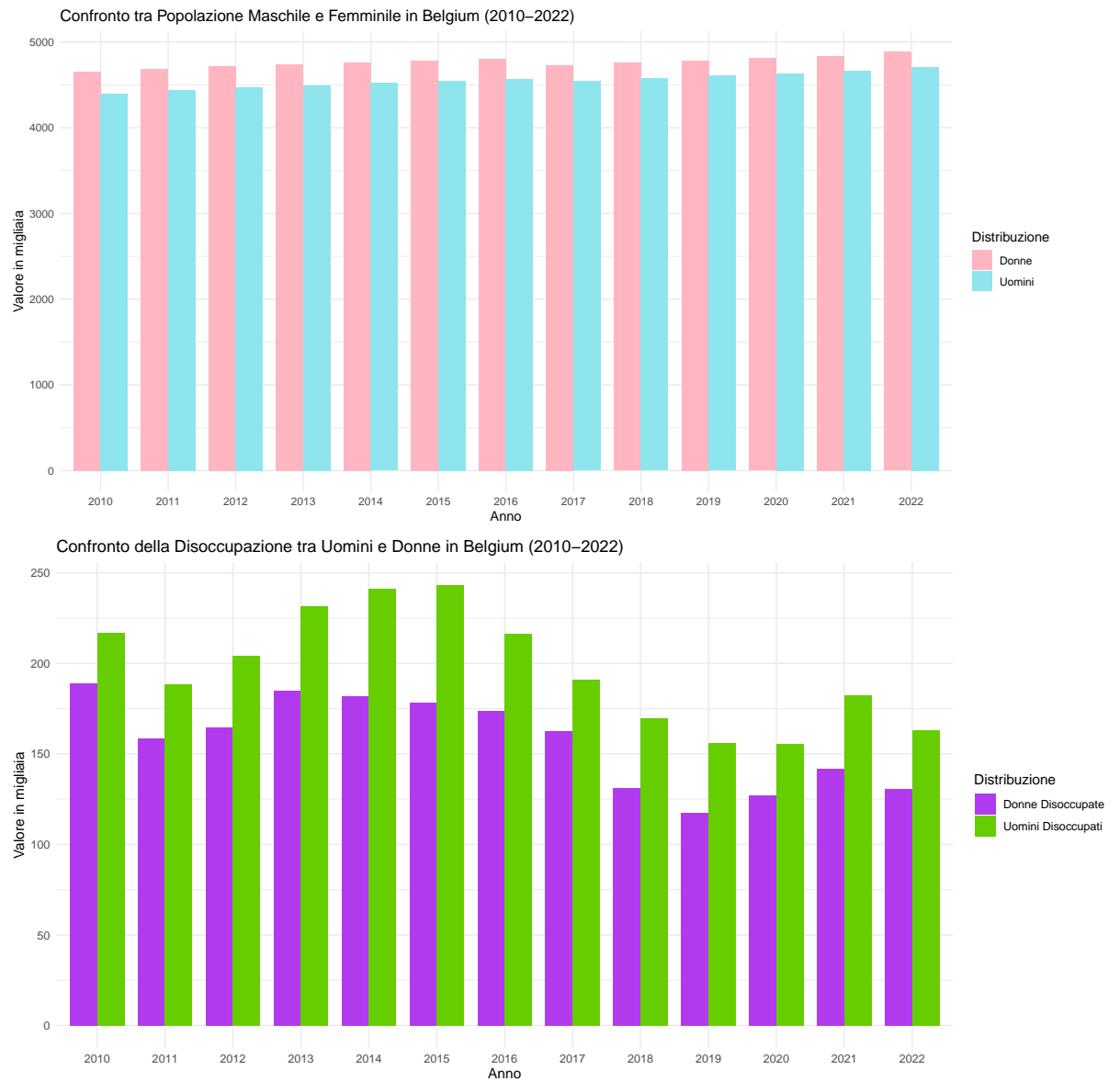
#### 3.1 Barplot sul sesso

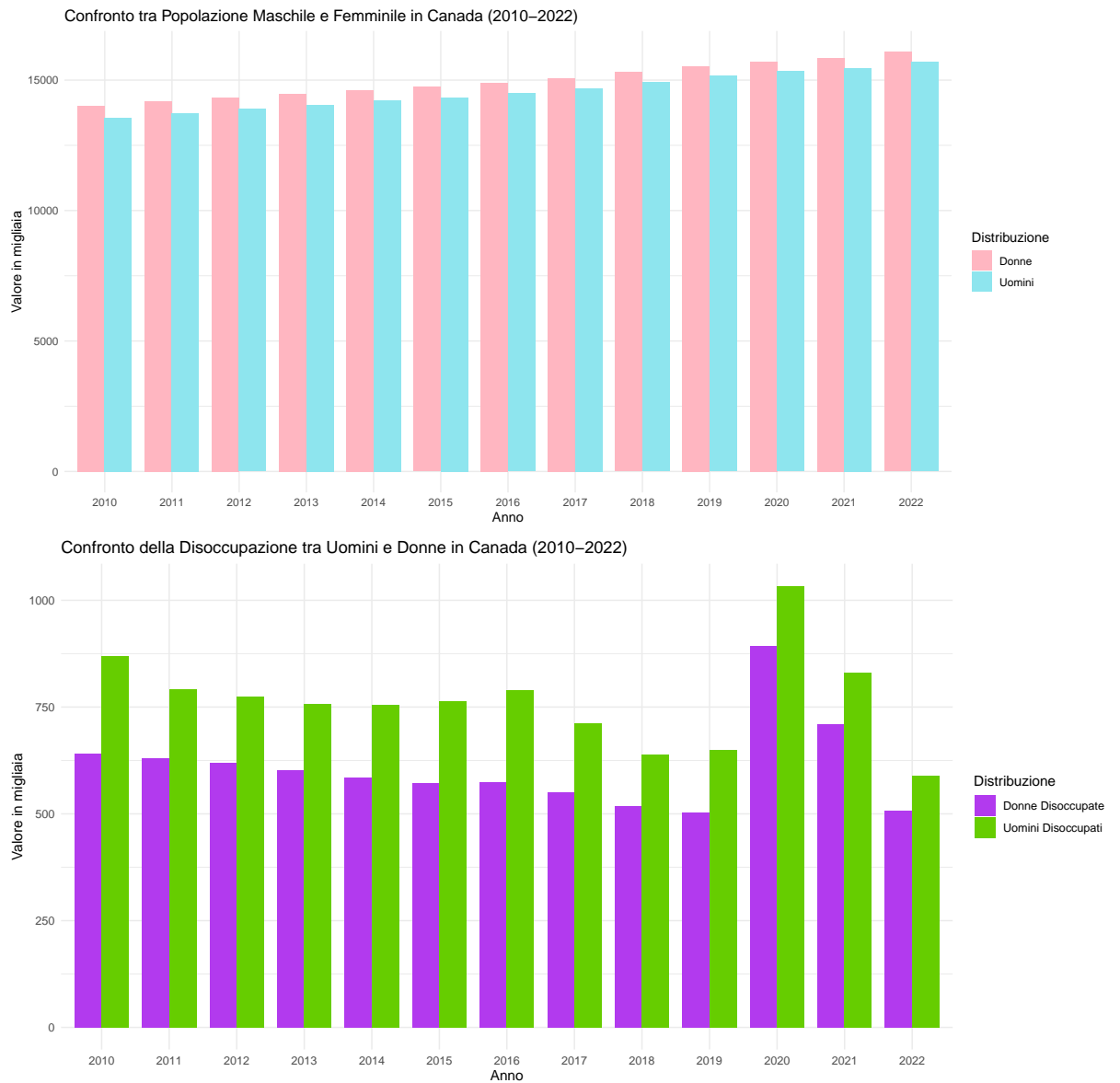
Per ogni paese vengono mostrati il grafico sulla distribuzione tra persone di sesso femminile e maschile all'interno di ogni paese, e la distribuzione di disoccupati di sesso maschile e femminile, dal 2010 al 2022. In questo modo è possibile fare un confronto sia tra popolazione maschile e femminile, sia tra disoccupazione maschile e femminile, sia tra popolazione e disoccupazione dello stesso sesso.

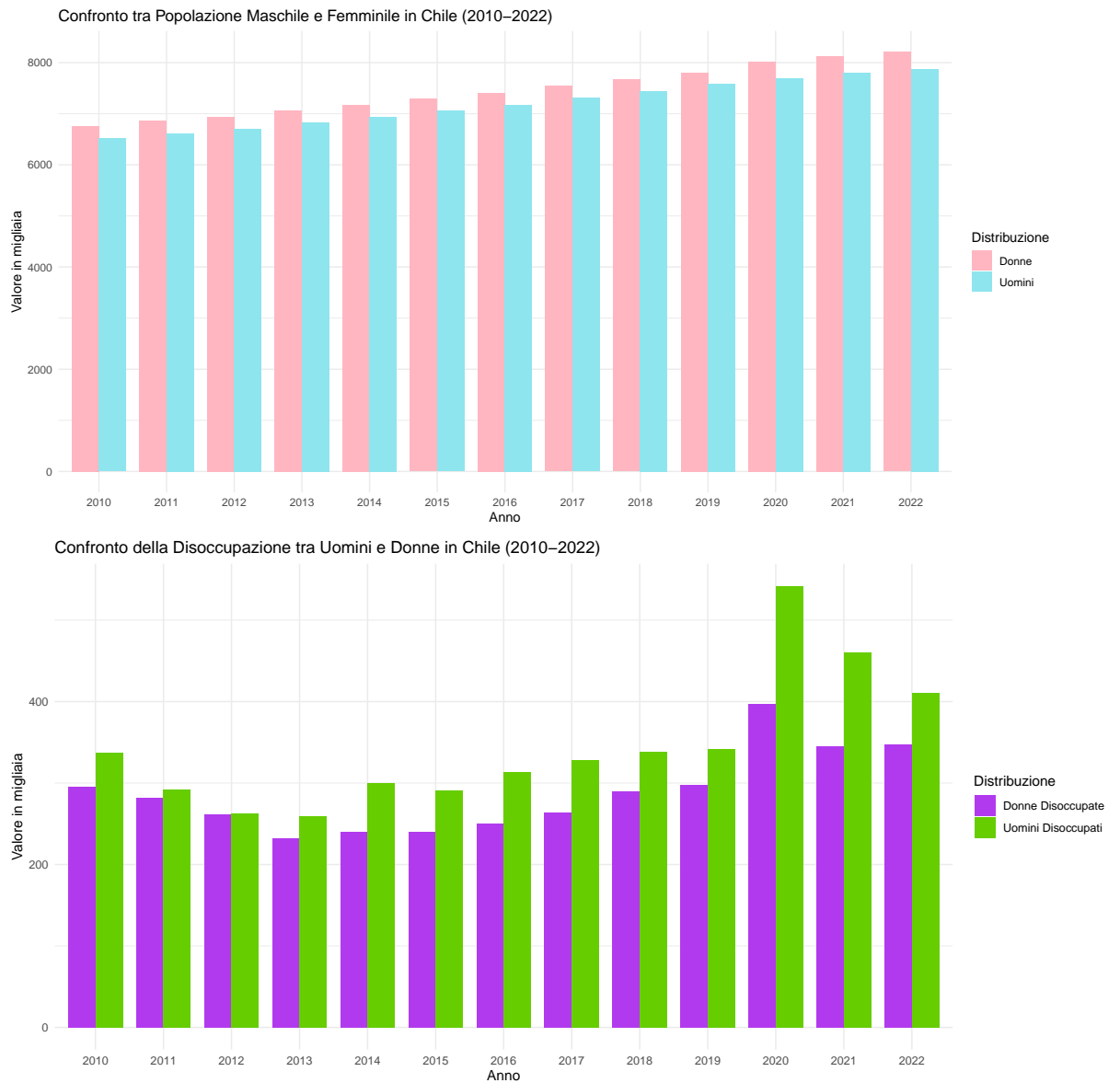


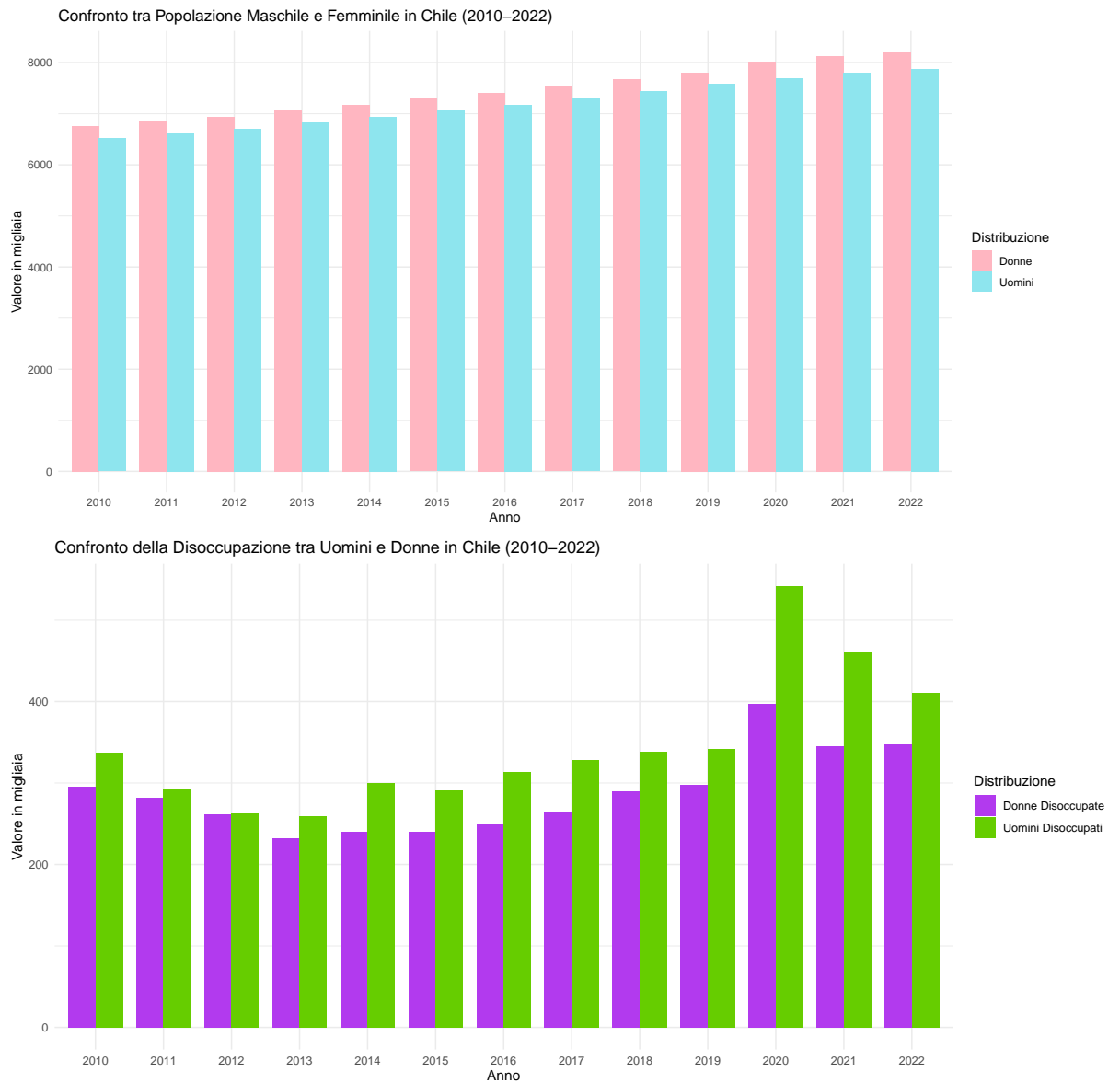


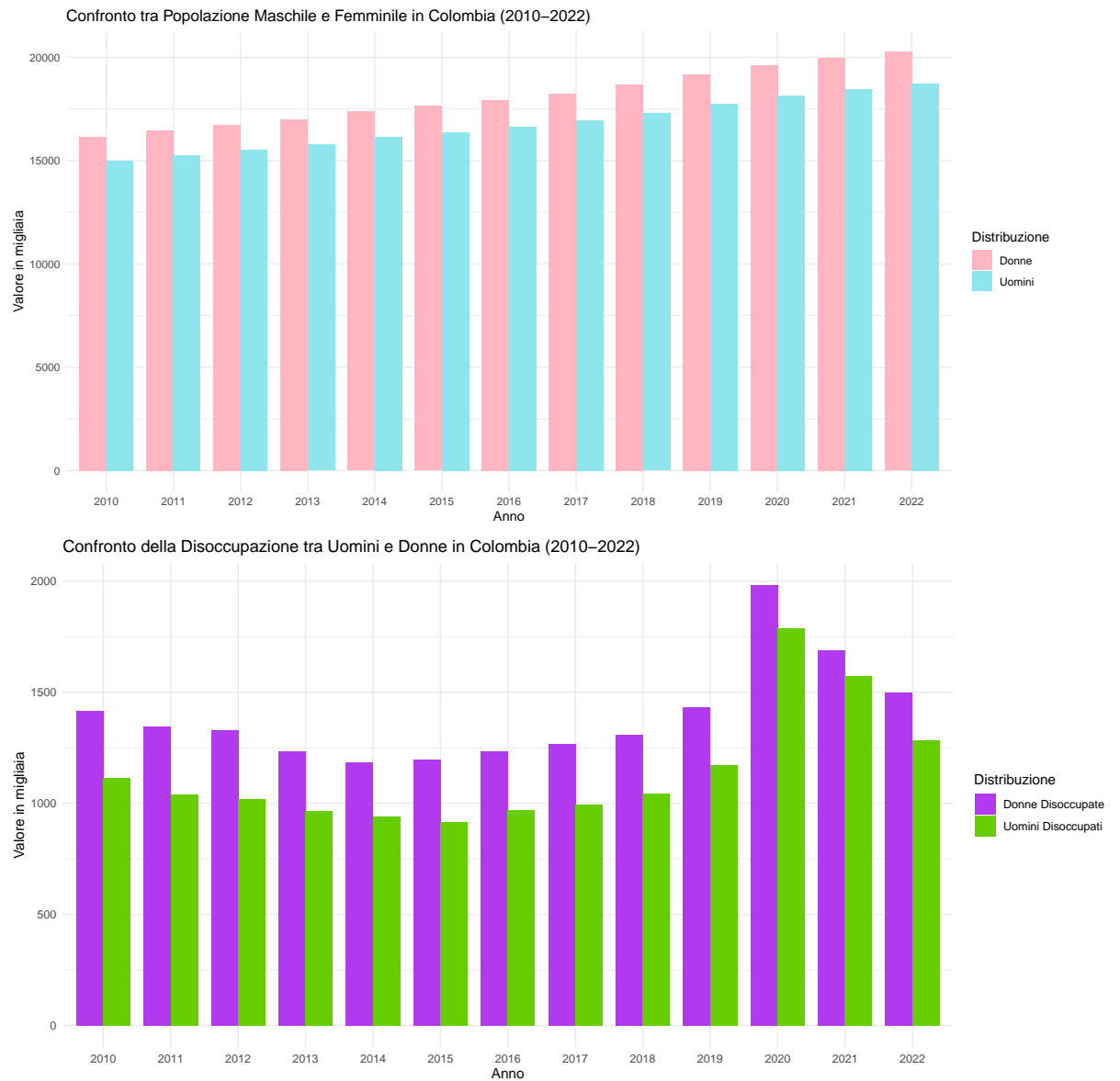


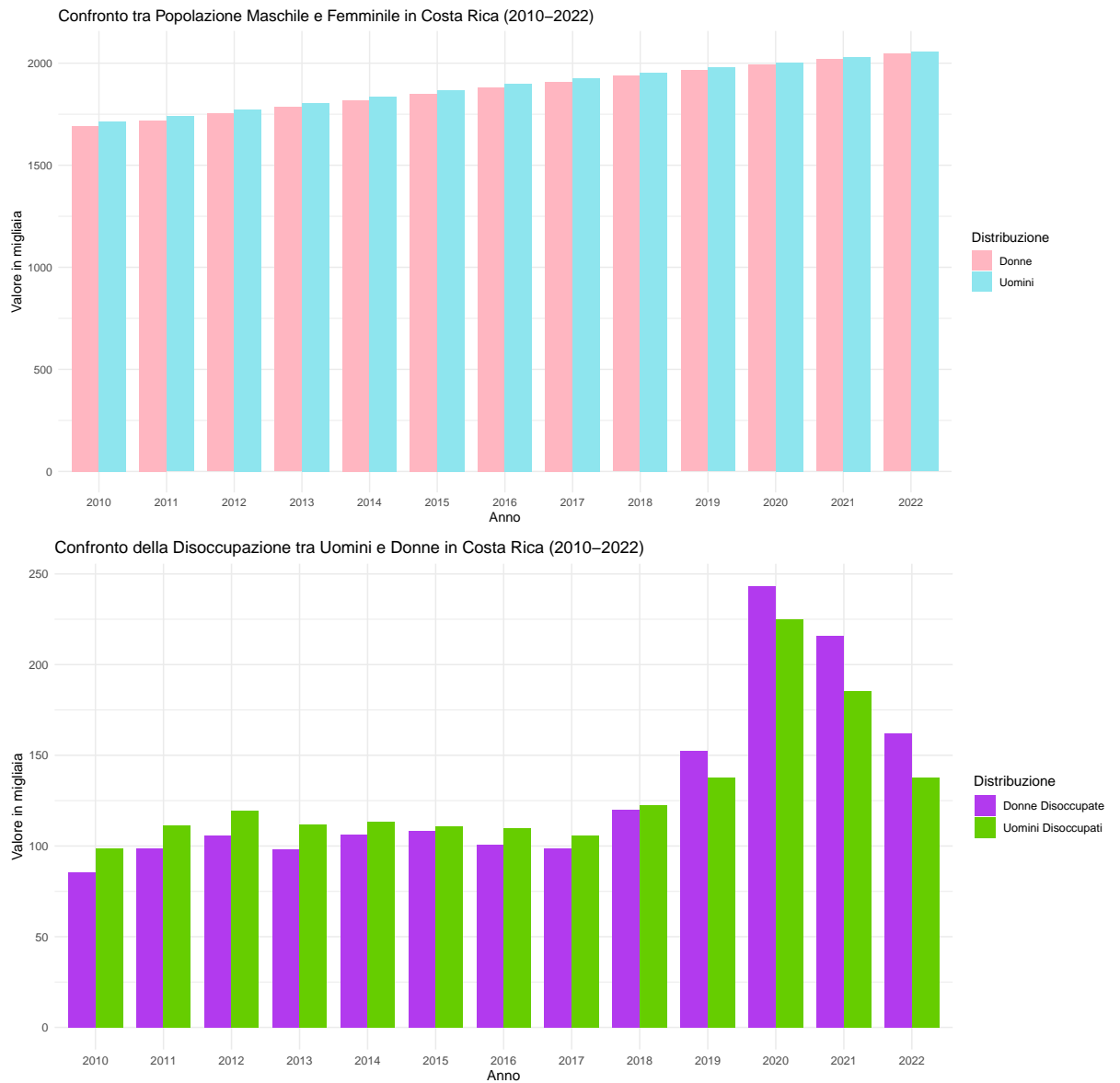






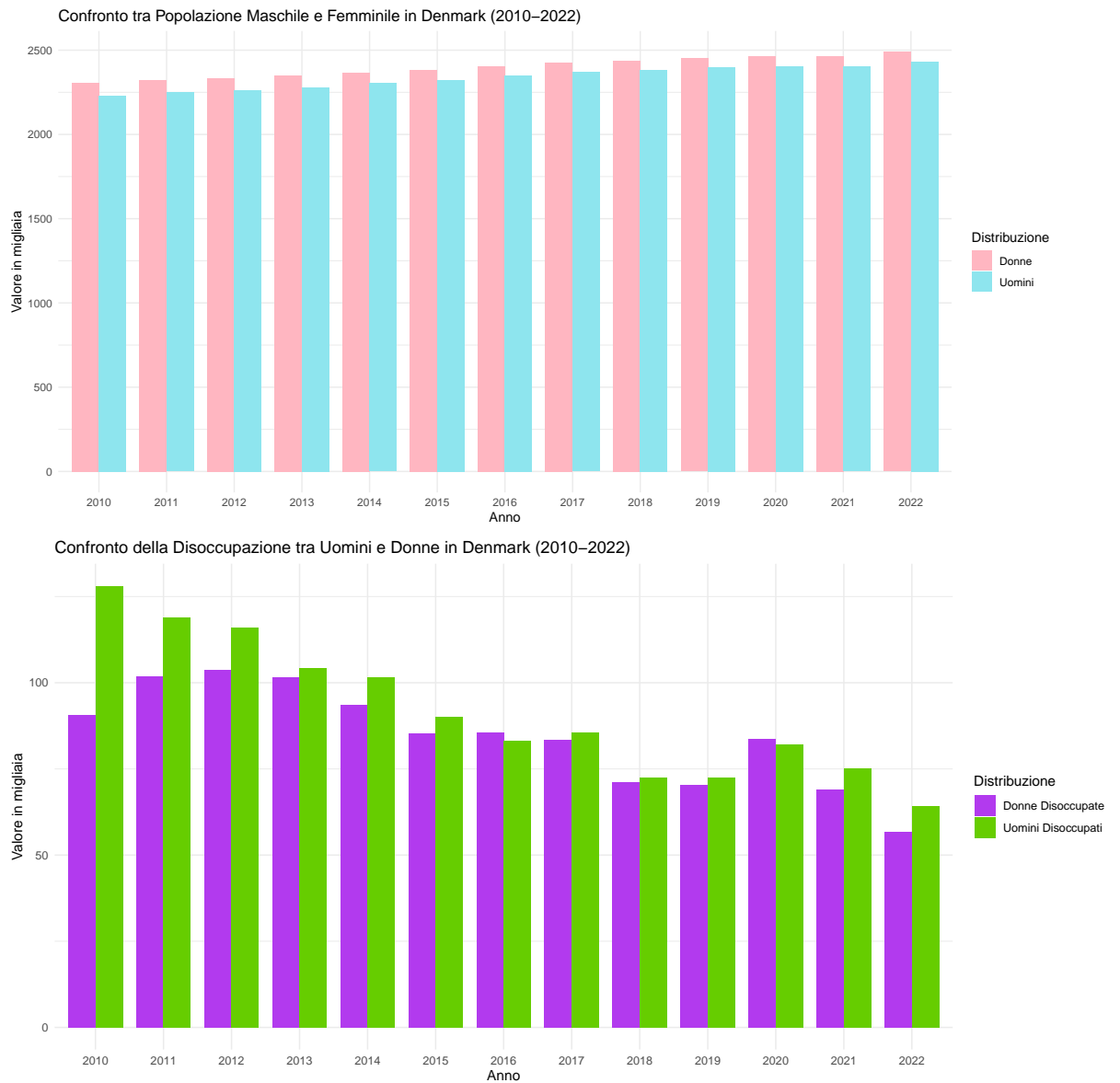






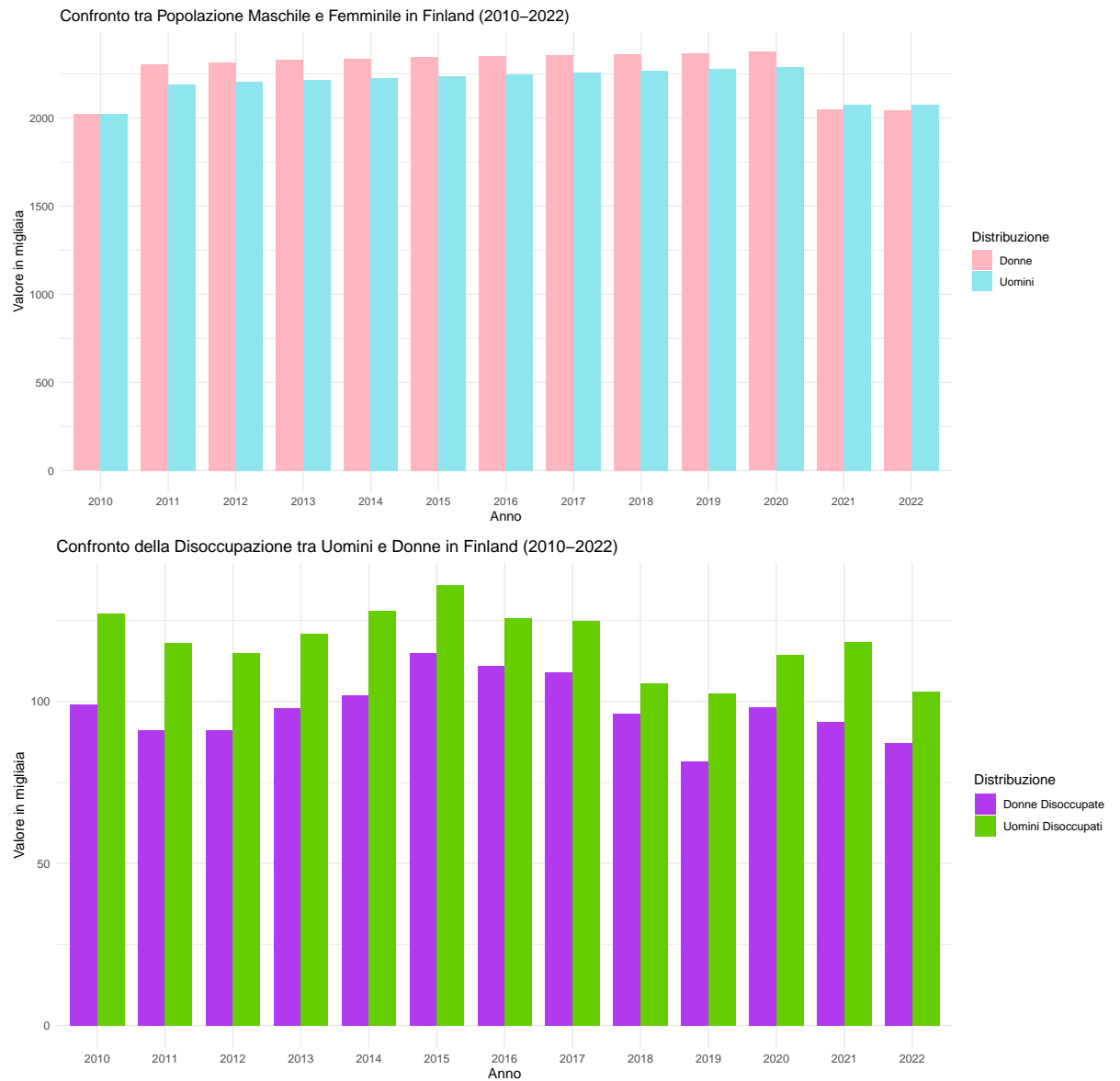


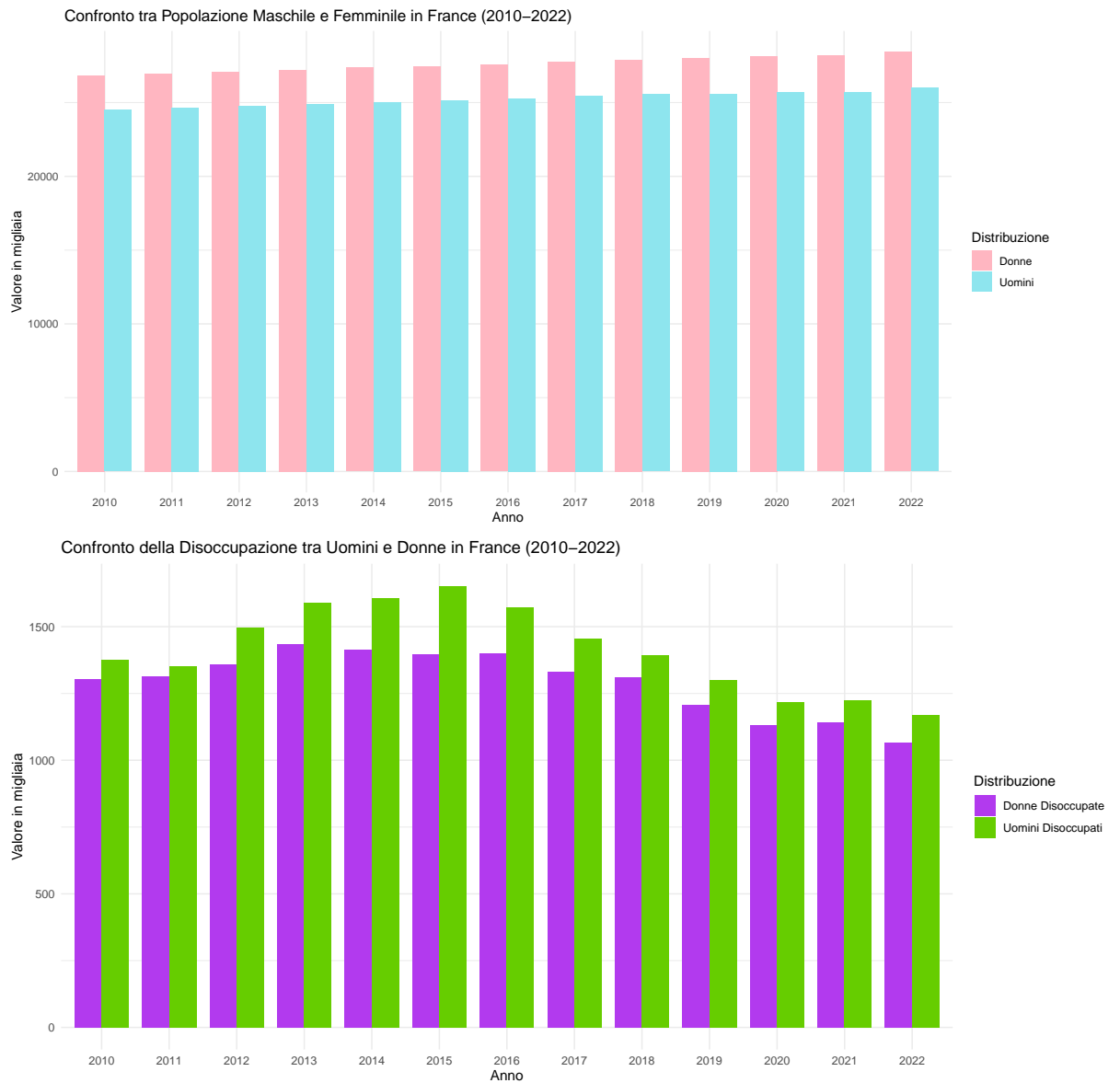




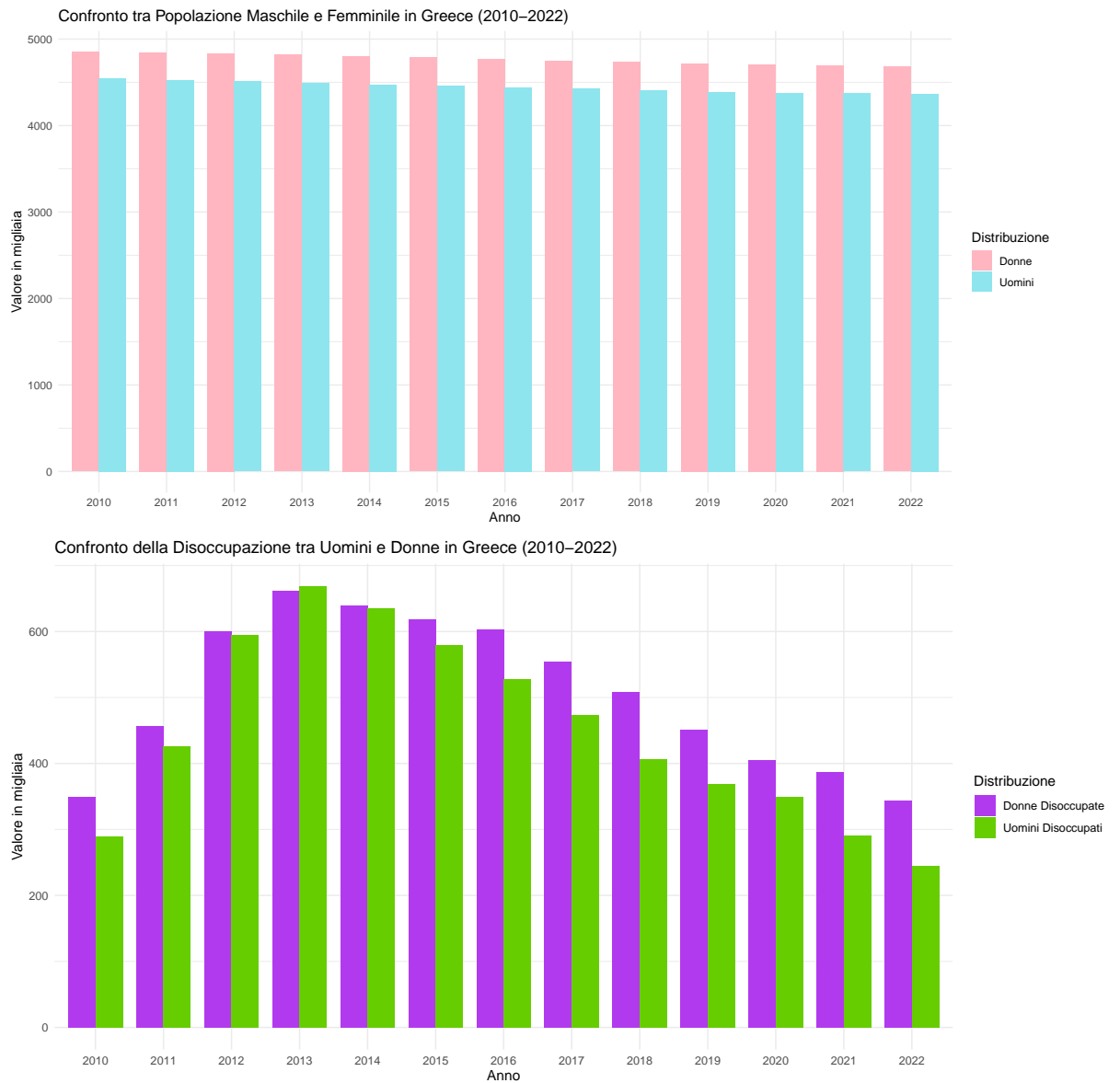


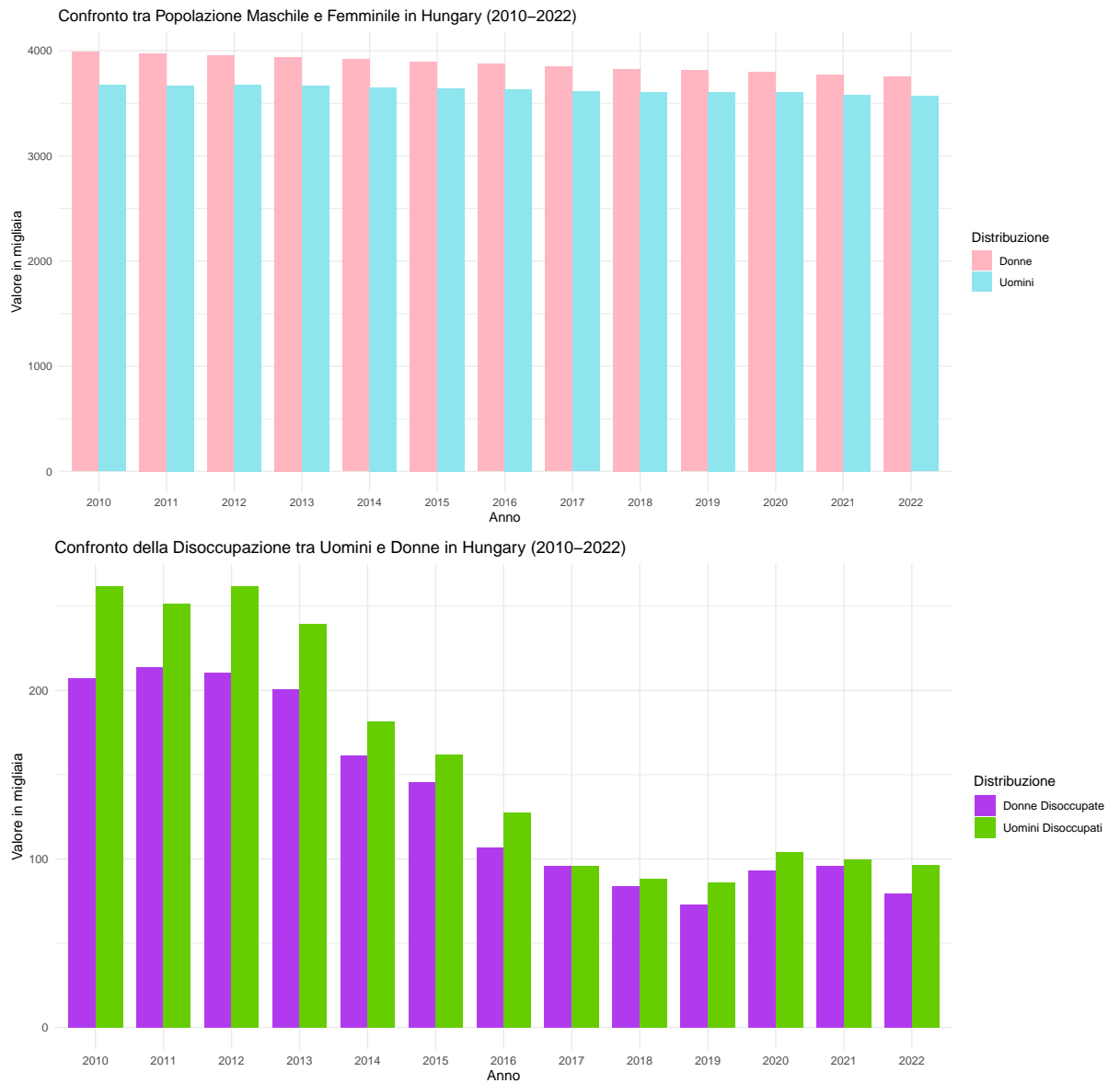
In Estonia il divario tra popolazione maschile e femminile è abbastanza considerevole. Potrebbe esserci un flusso migratorio che porta più uomini ad emigrare dall'Estonia per lavoro o altre opportunità, mentre le donne potrebbero rimanere nel paese. Questo potrebbe essere il motivo della maggioranza femminile nella popolazione [12]. Nonostante la popolazione femminile sia superiore, la disoccupazione femminile è inferiore a quella maschile. Ciò potrebbe essere dovuto al fatto che lavori generalmente considerati femminili (come ad esempio l'infermiera) richiedano meno qualificazioni.



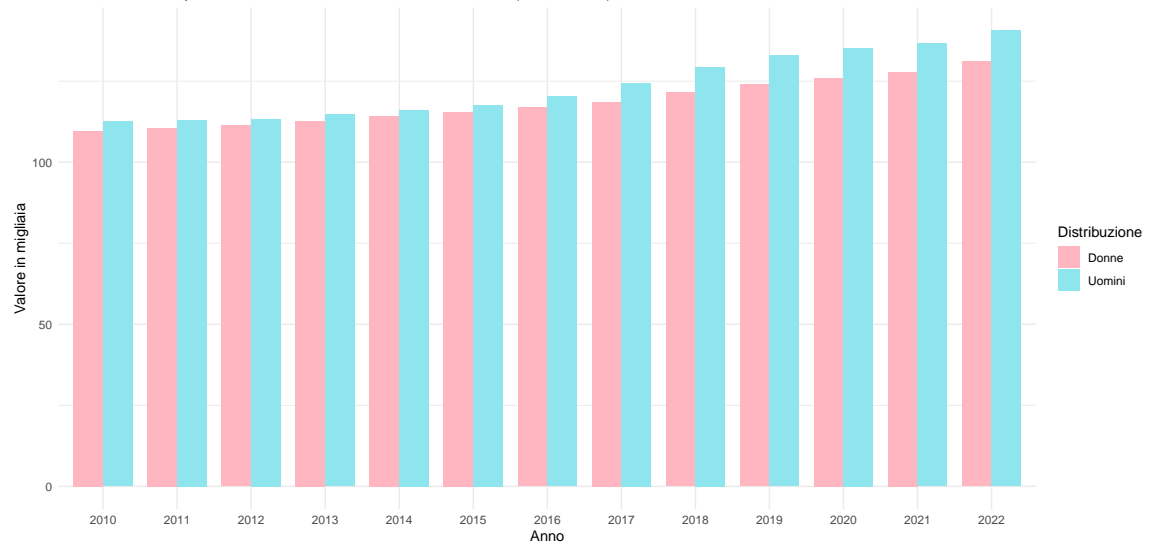




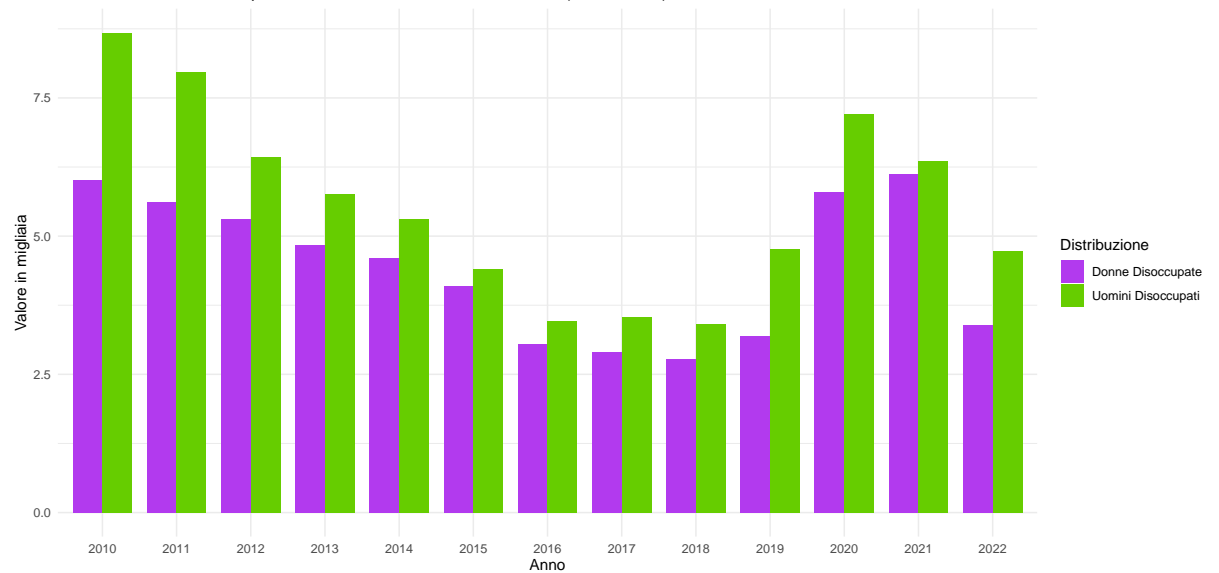




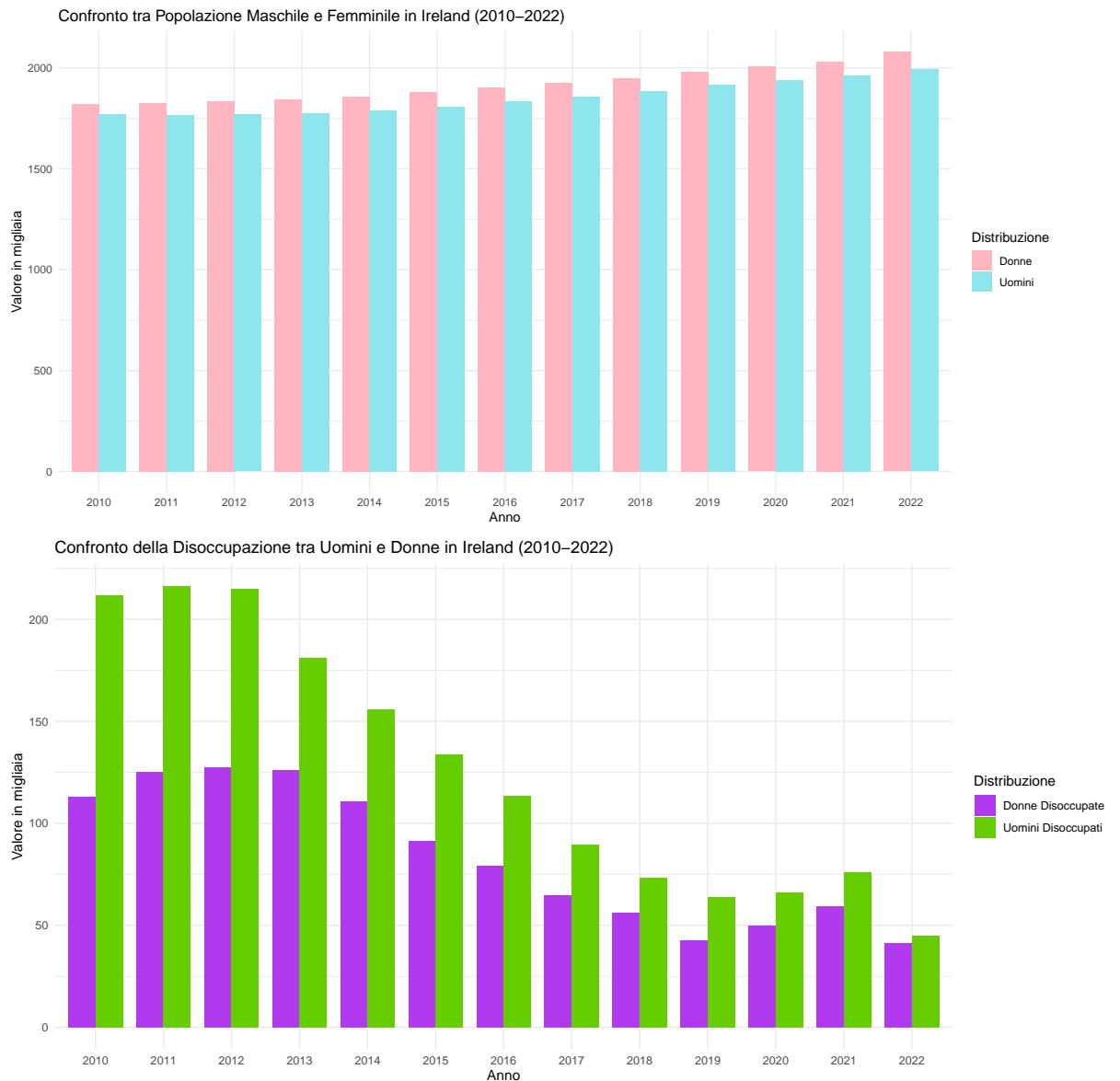
Confronto tra Popolazione Maschile e Femminile in Iceland (2010–2022)



Confronto della Disoccupazione tra Uomini e Donne in Iceland (2010–2022)

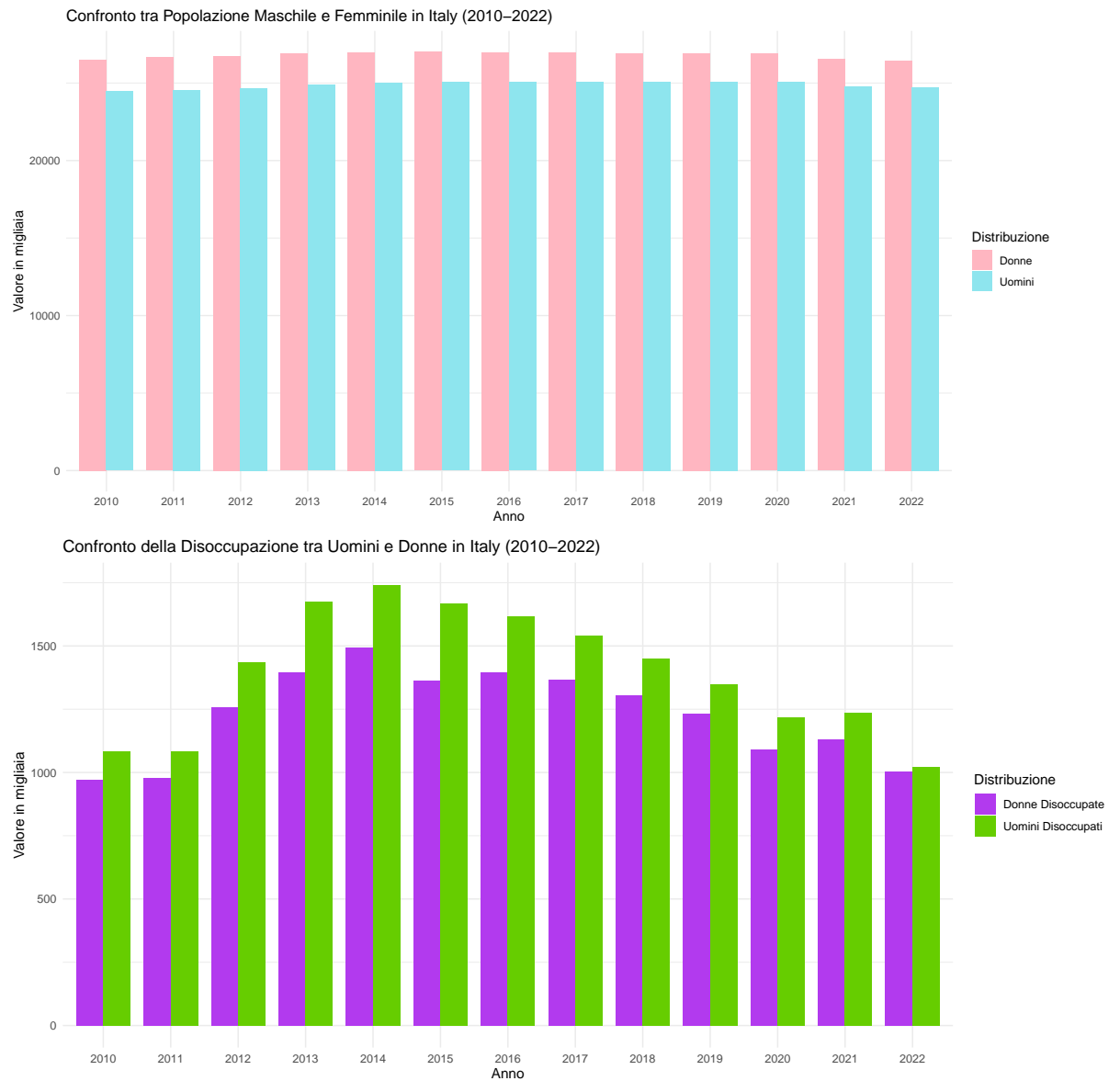


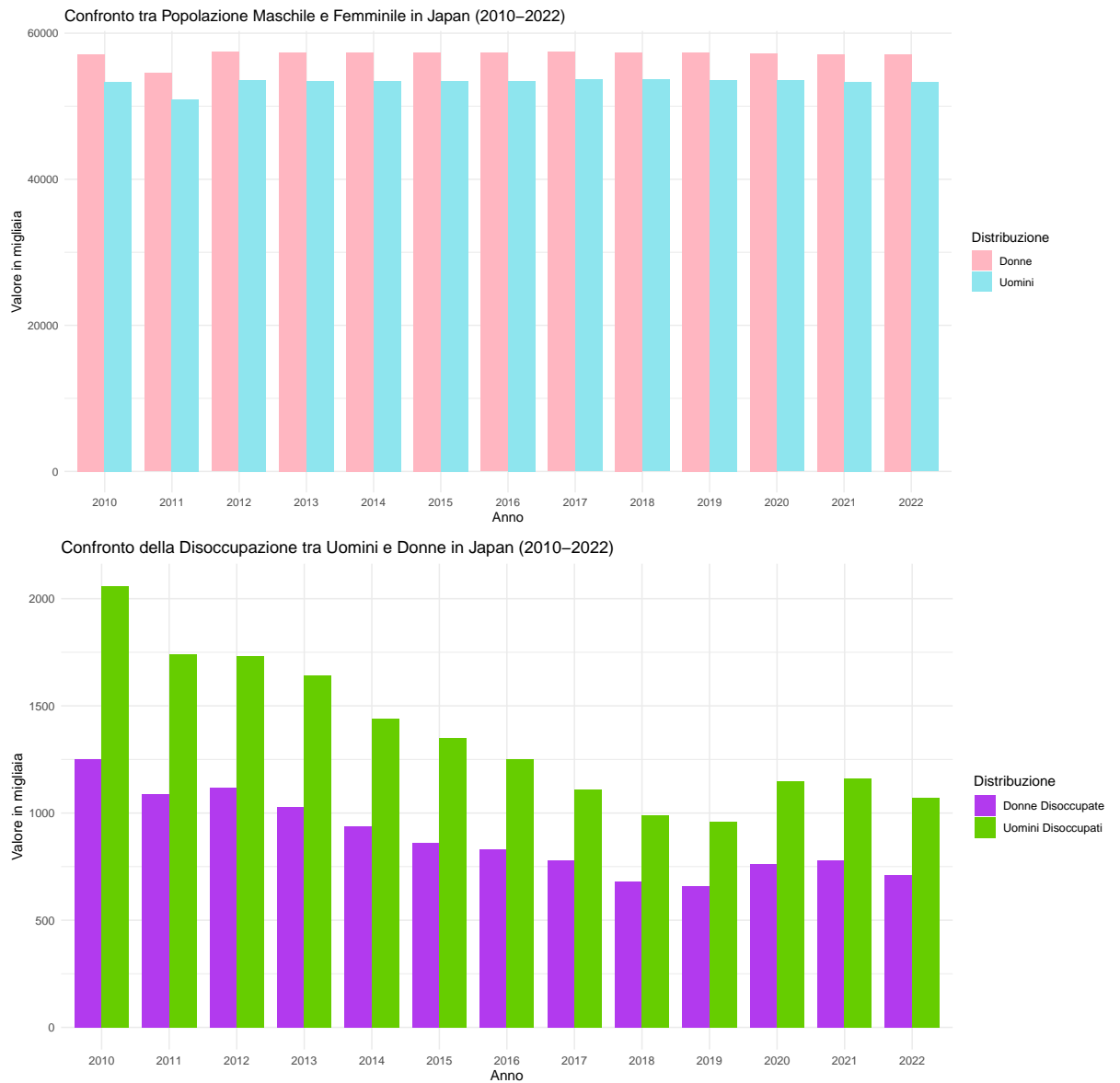


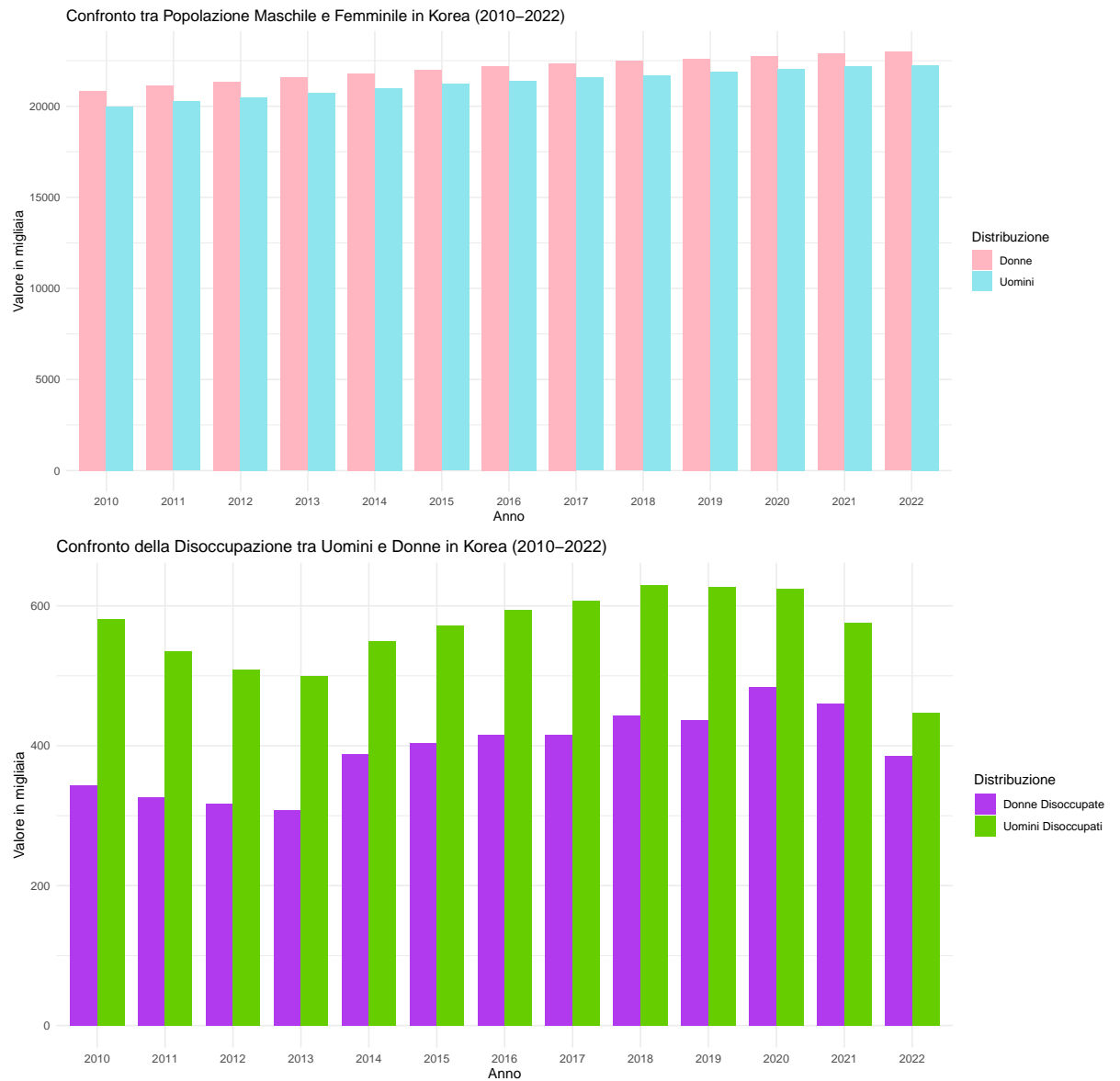


L'Irlanda, il Messico, la Turchia, la Germania, Il Giappone, la Korea e la Lituania hanno una caratteristica in comune: la percentuale di disoccupati maschile, nella maggior parte degli anni, è quasi il doppio di quella femminile. Due motivi potrebbero essere i seguenti: In alcuni paesi, potrebbe esserci una segregazione occupazionale basata sul genere, con gli uomini più concentrati in settori che sono stati particolarmente colpiti dalla disoccupazione, come il settore manifatturiero o edile; Gli uomini potrebbero avere preferenze lavorative diverse dalle donne, scegliendo settori o posizioni che sono più vulnerabili alla disoccupazione o più influenzati dalle fluttuazioni economiche.





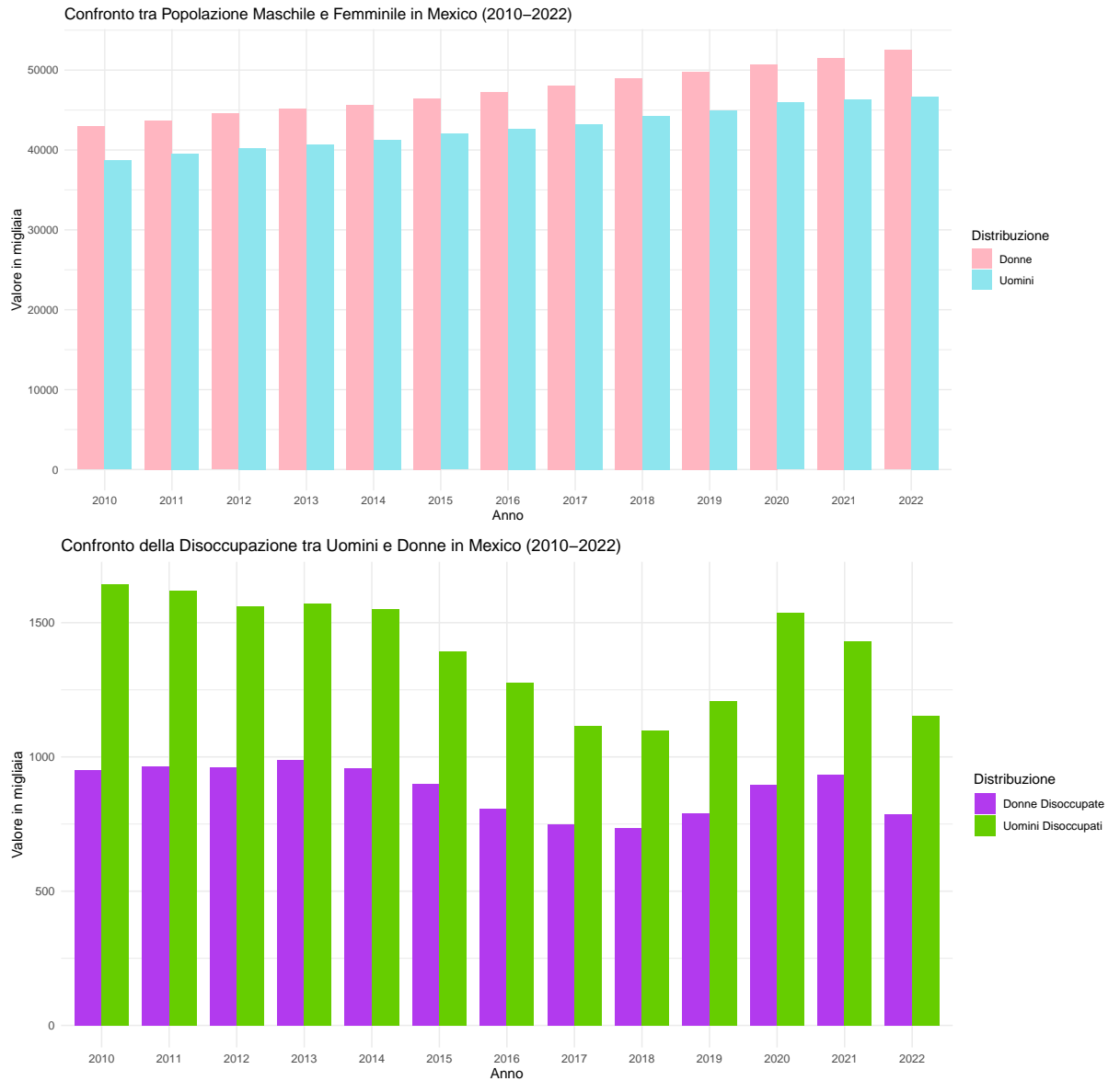






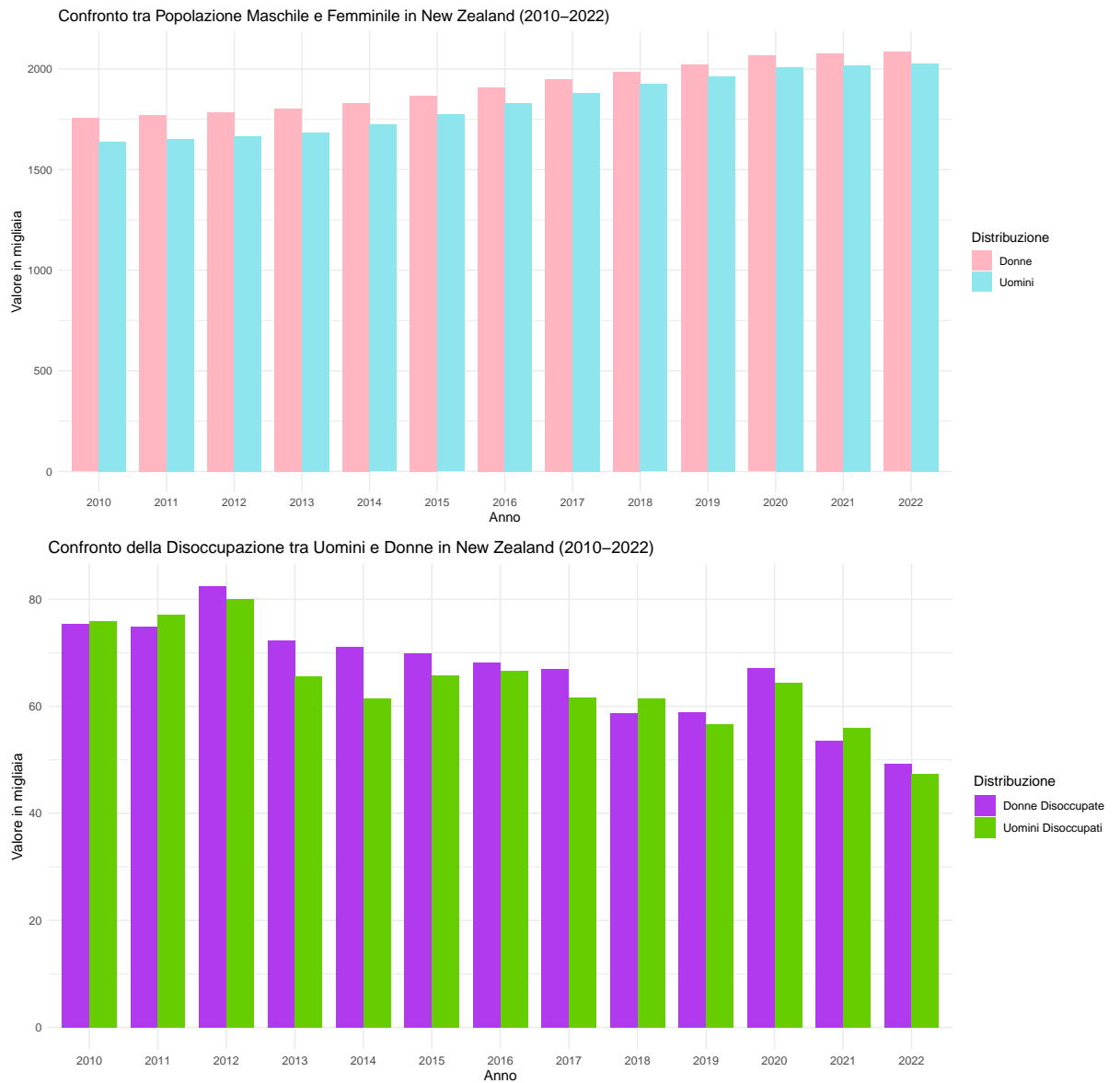
La Lituania è l'unico tra i 38 paesi ad avere nella popolazione molte più persone di sesso maschile che femminile. È possibile che una percentuale significativa di donne lituane abbia scelto di emigrare in cerca di opportunità di lavoro o migliori condizioni di vita in altri paesi. Infatti la Lituania è tra i paesi con la più alta "gender inequality" [13].

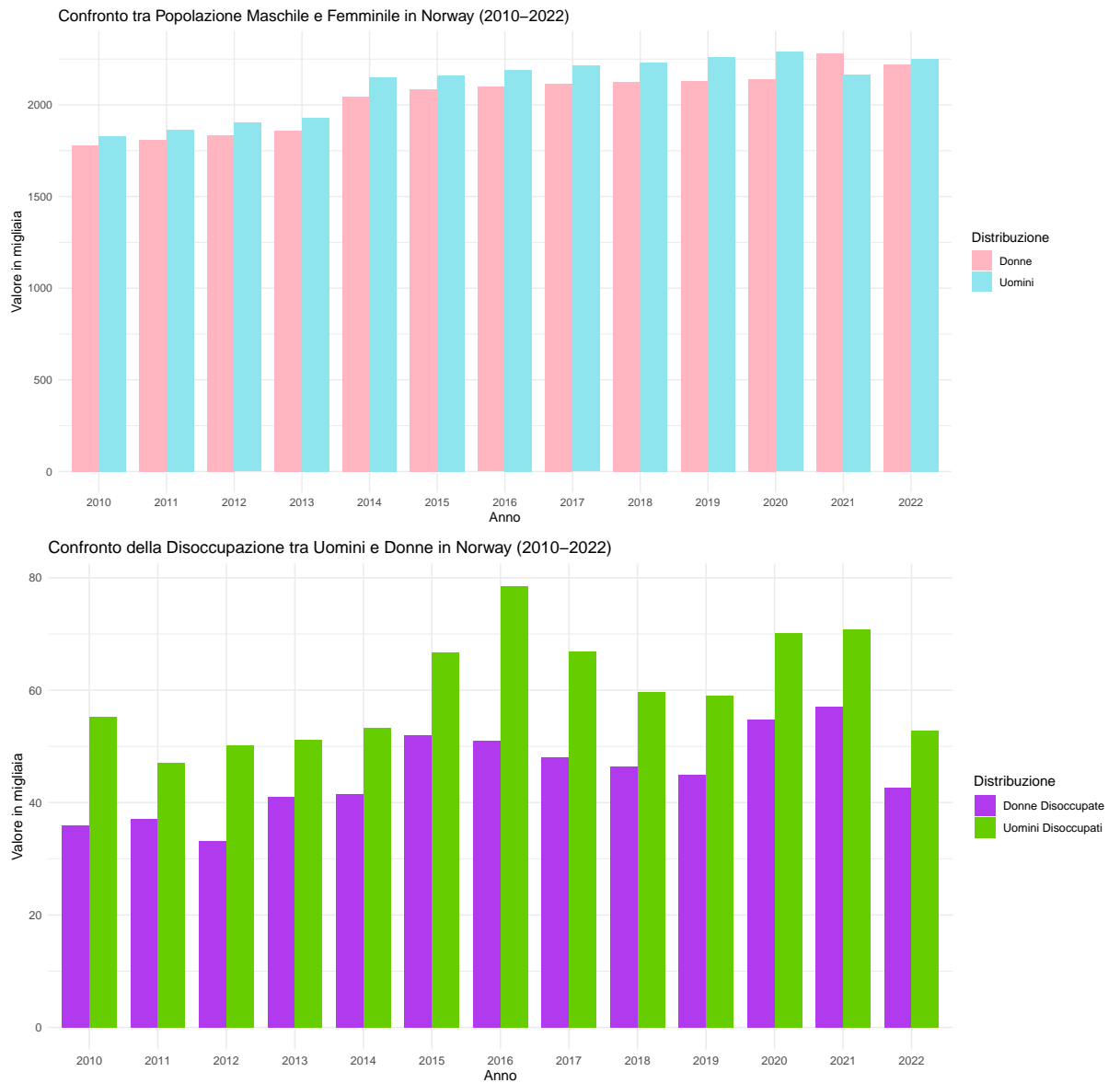


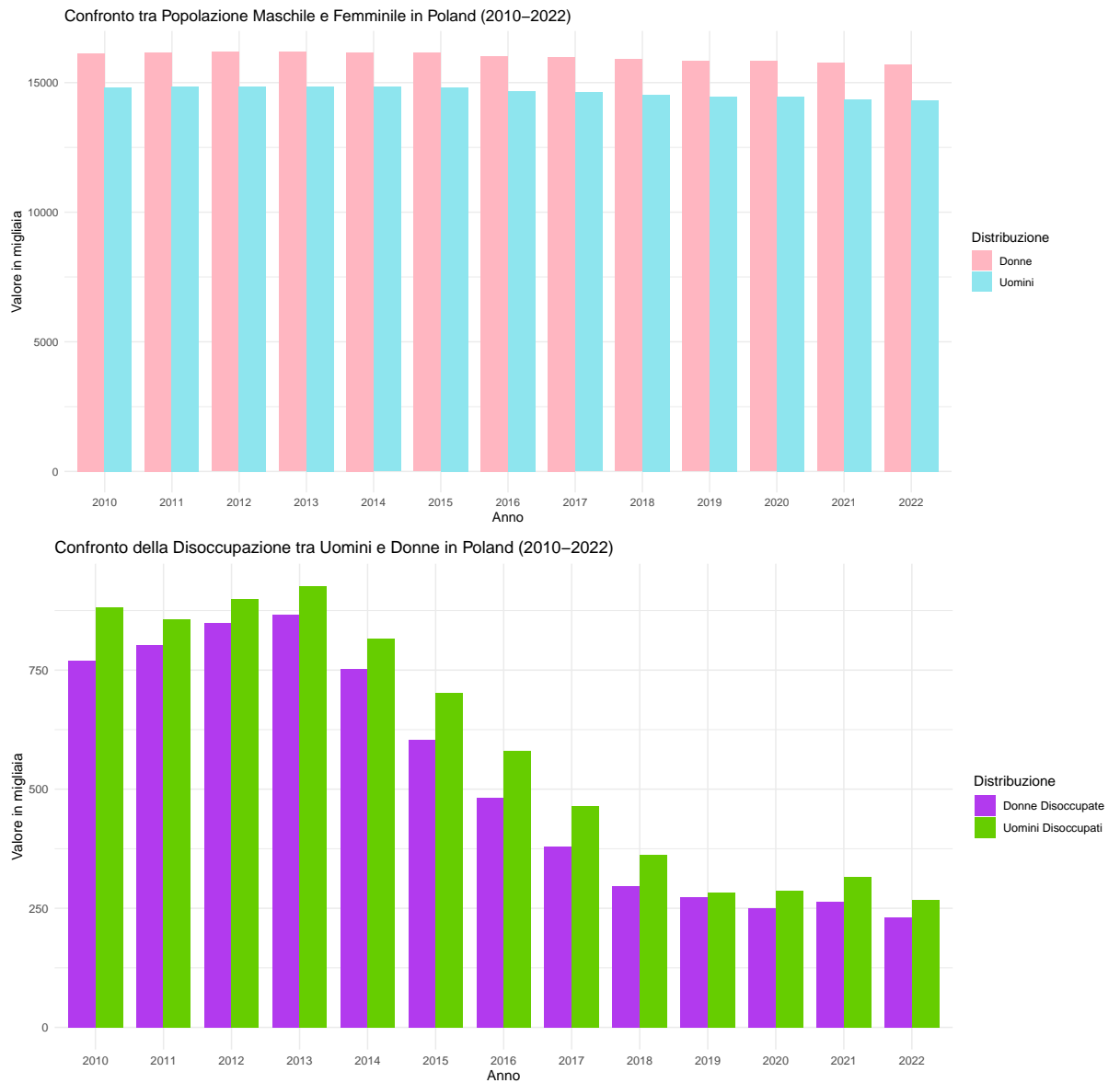






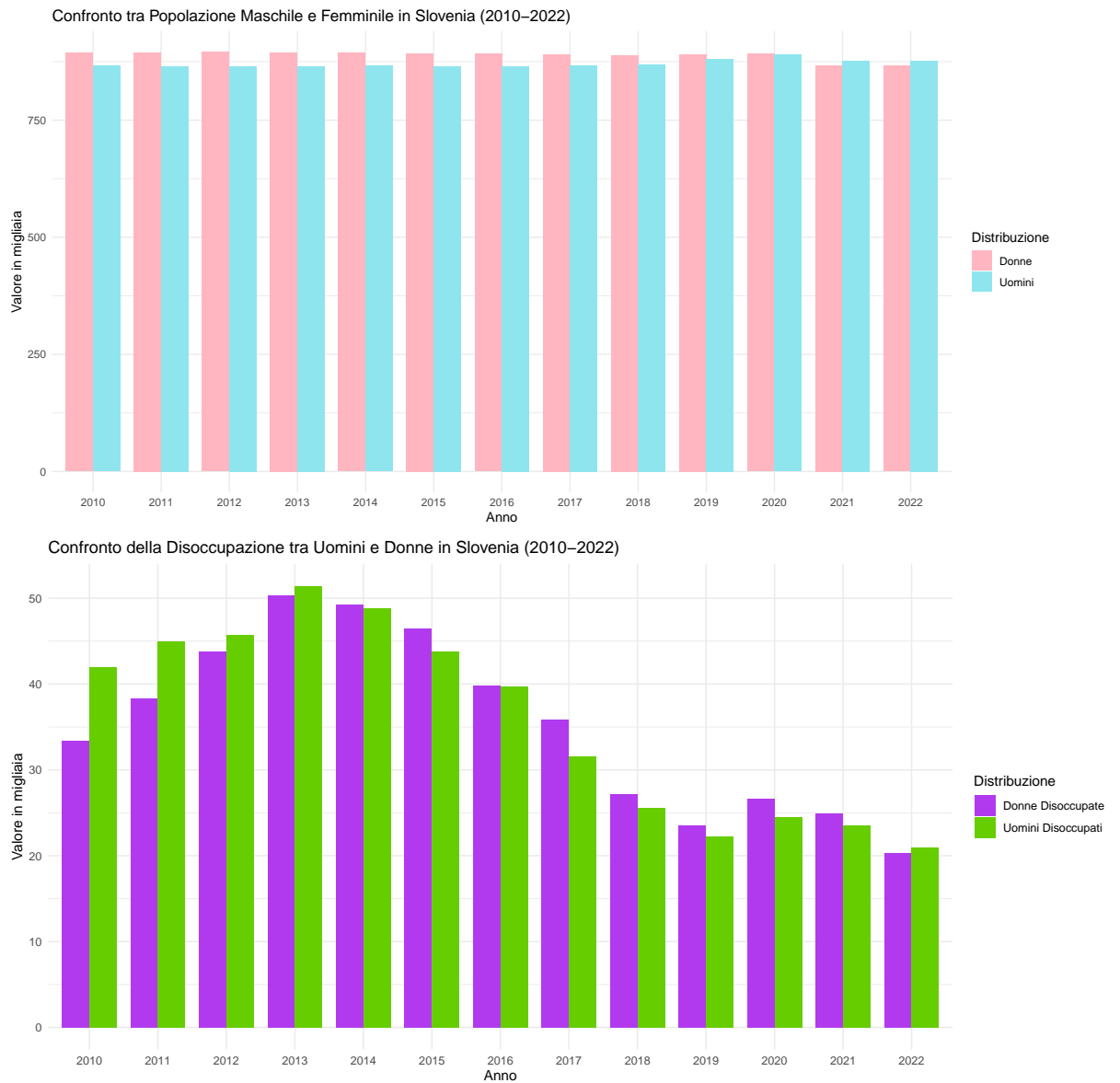


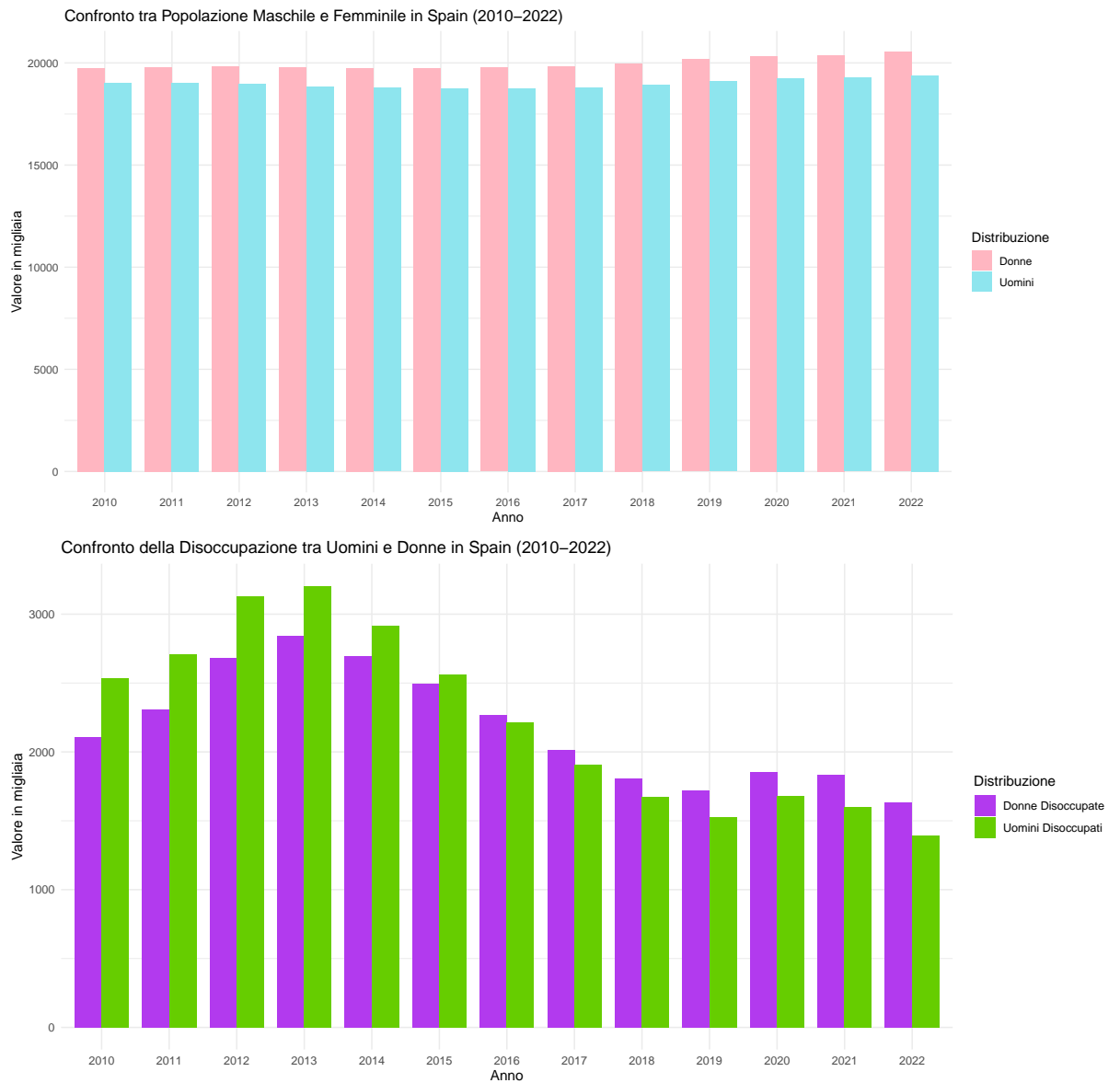




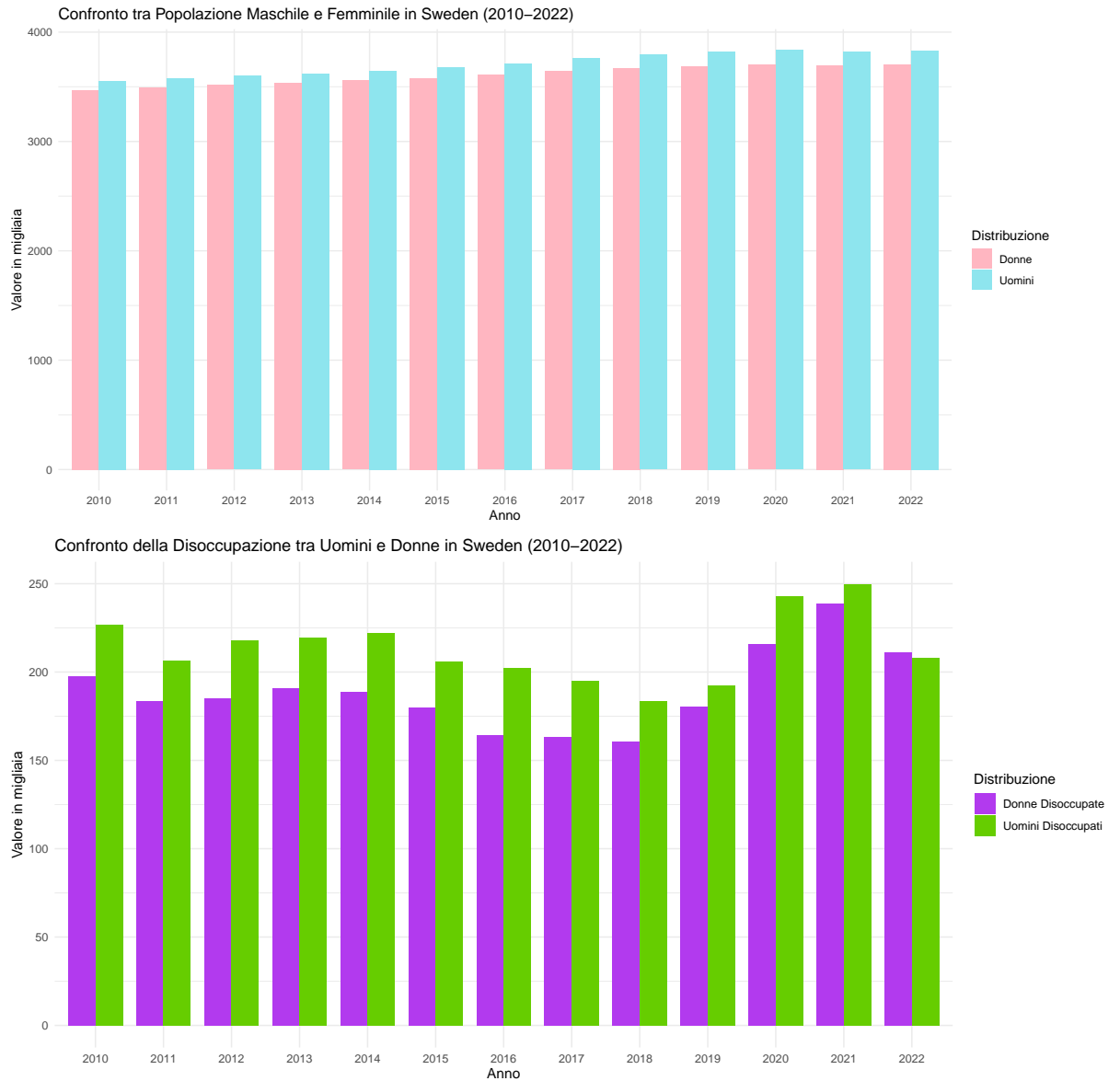


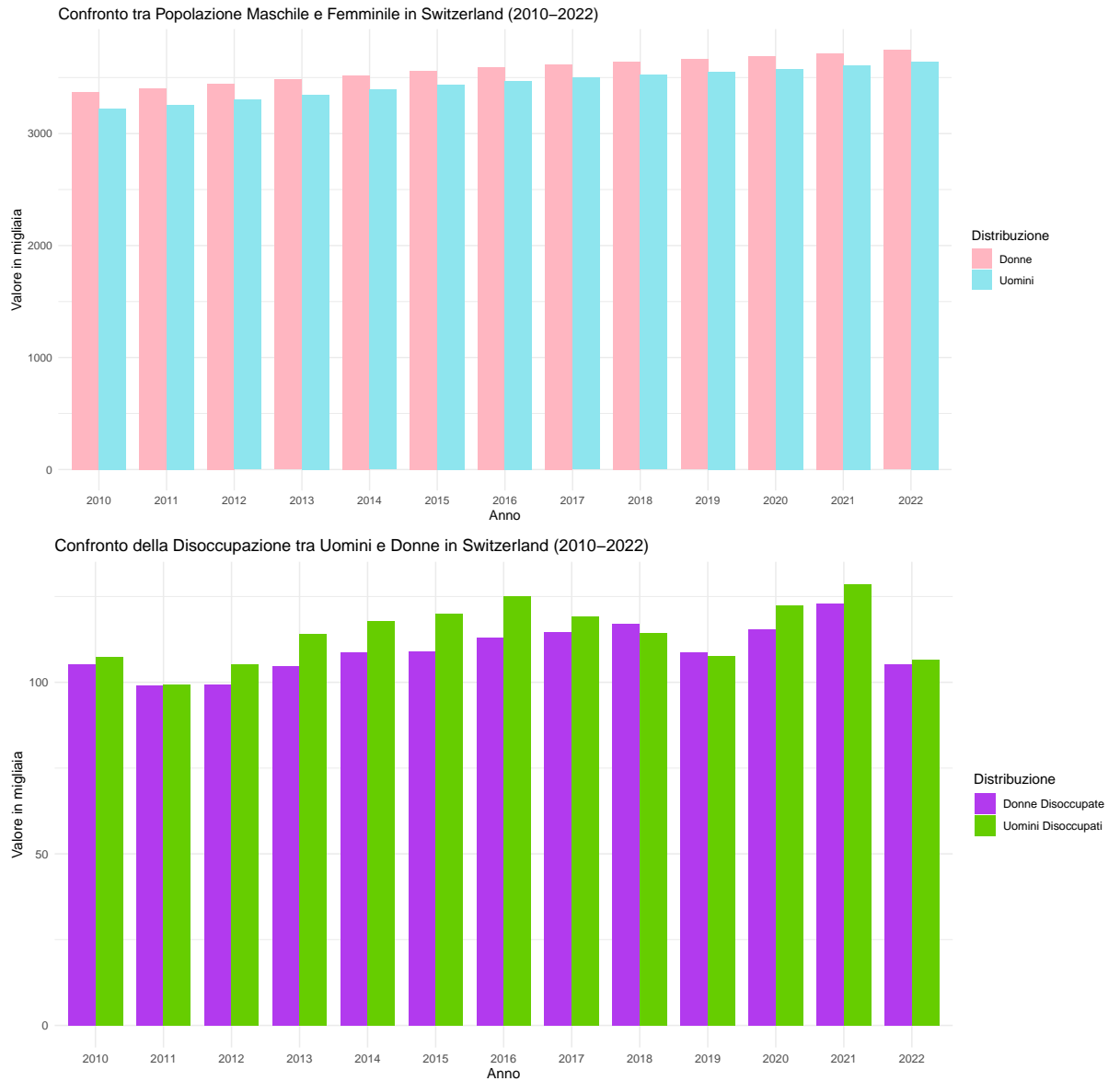


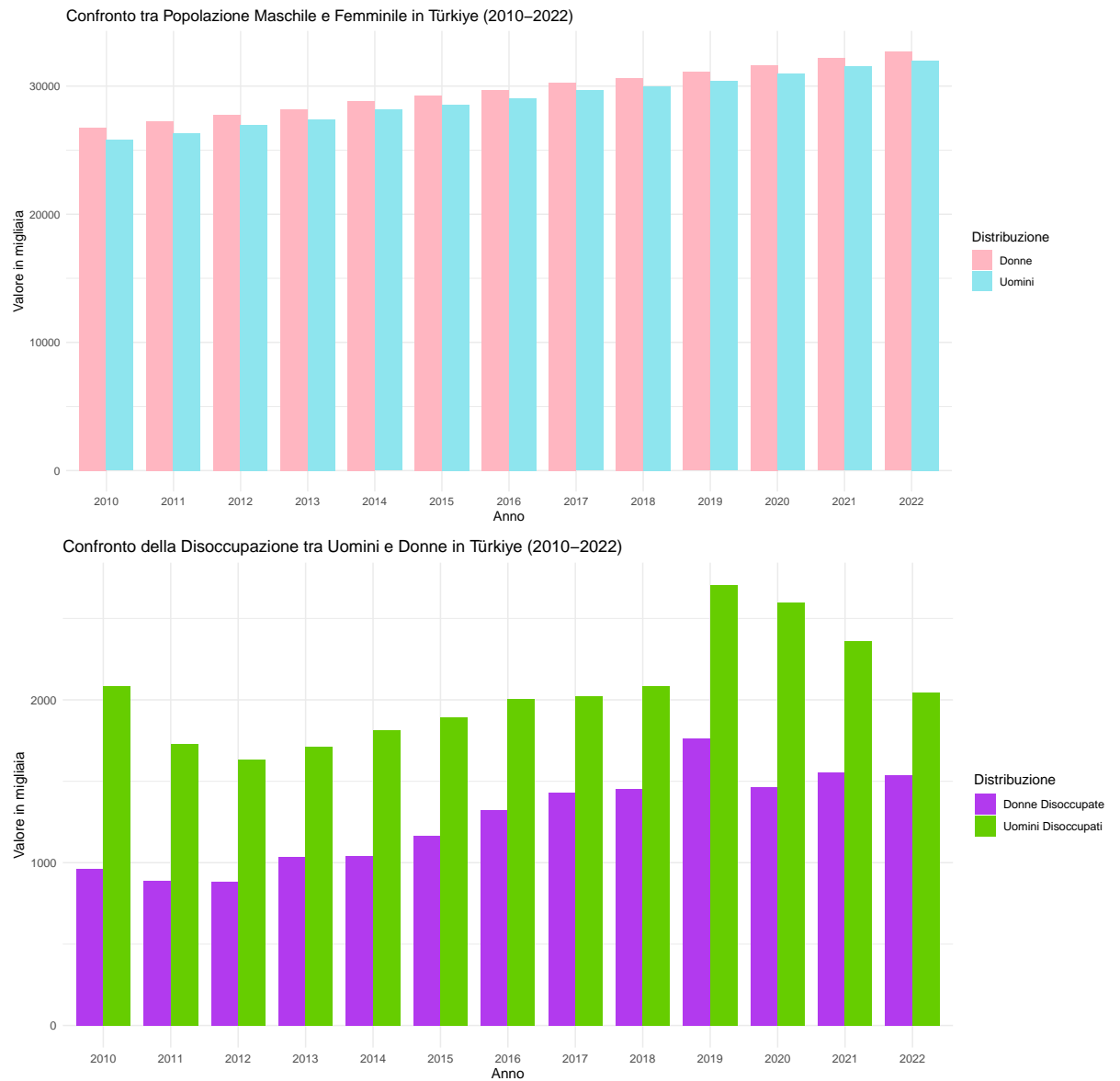






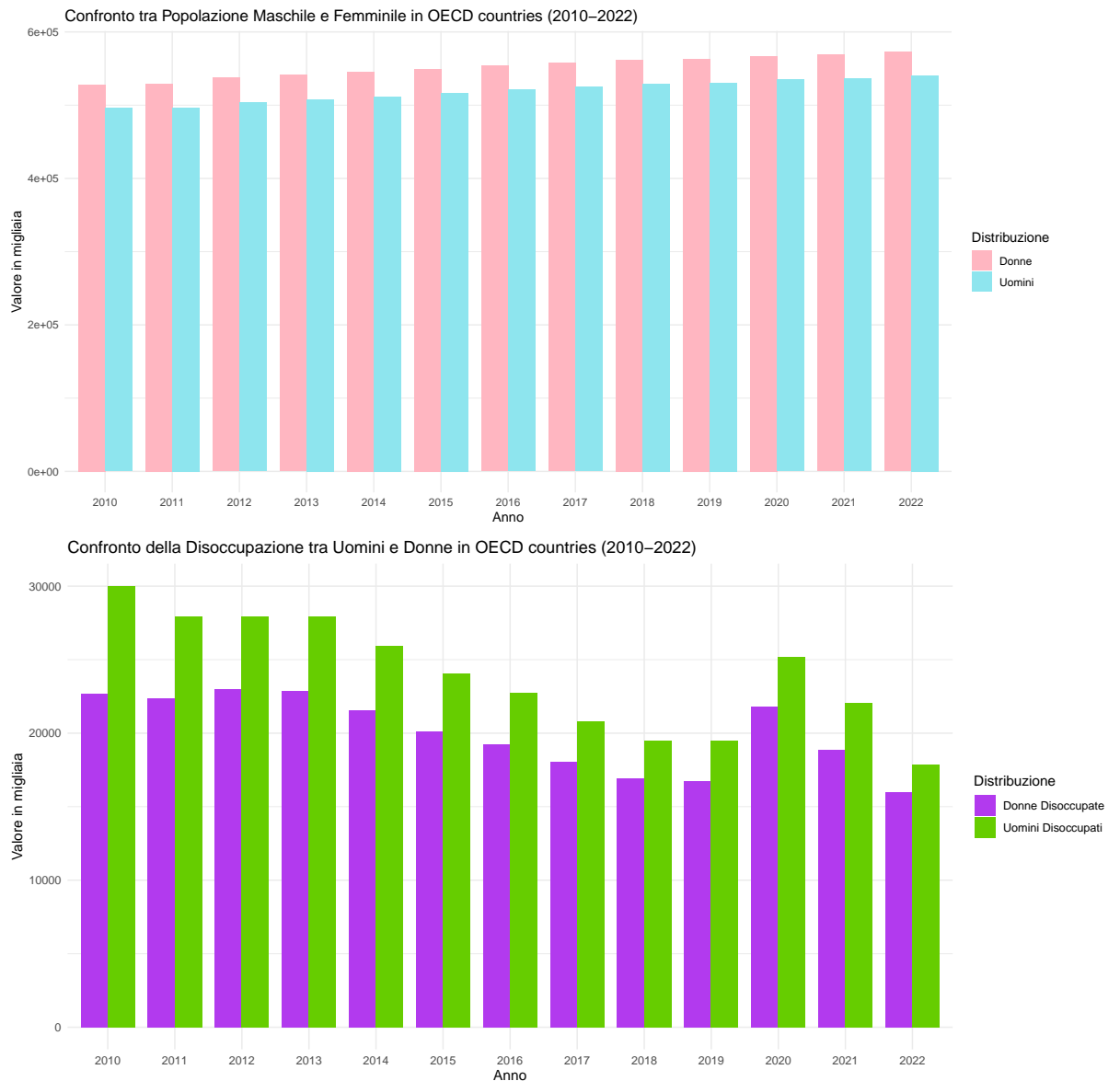












In media, la popolazione femminile in ogni paese tende ad essere maggiore di quella maschile. Nonostante ciò, si osserva spesso che la disoccupazione degli uomini è più elevata rispetto a quella delle donne.

Questo fenomeno può essere attribuito a diversi fattori interconnessi. In primo luogo, i ruoli di genere (o *stereotipi di genere*? [14]) e stereotipi culturali svolgono un ruolo significativo nella determinazione delle aspettative lavorative e delle scelte di carriera sia per gli uomini che per le donne. Spesso si associa la figura femminile al ruolo di casalinga, preoccupata principalmente delle faccende domestiche e della cura dei figli, mentre si presume che l'uomo sia l'unico a sostenere finanziariamente la famiglia attraverso il lavoro retribuito.

Nonostante l'importanza e l'impegno richiesto dalle casalinghe, ciò non viene riconosciuto come lavoro nel senso tradizionale e non comporta compensazioni economiche o benefit aziendali come le ferie pagate o le malattie concesse. Questa mancanza di riconoscimento formale può influenzare la distribuzione del lavoro retribuito tra i sessi e contribuire alla disparità di disoccupazione. [15]

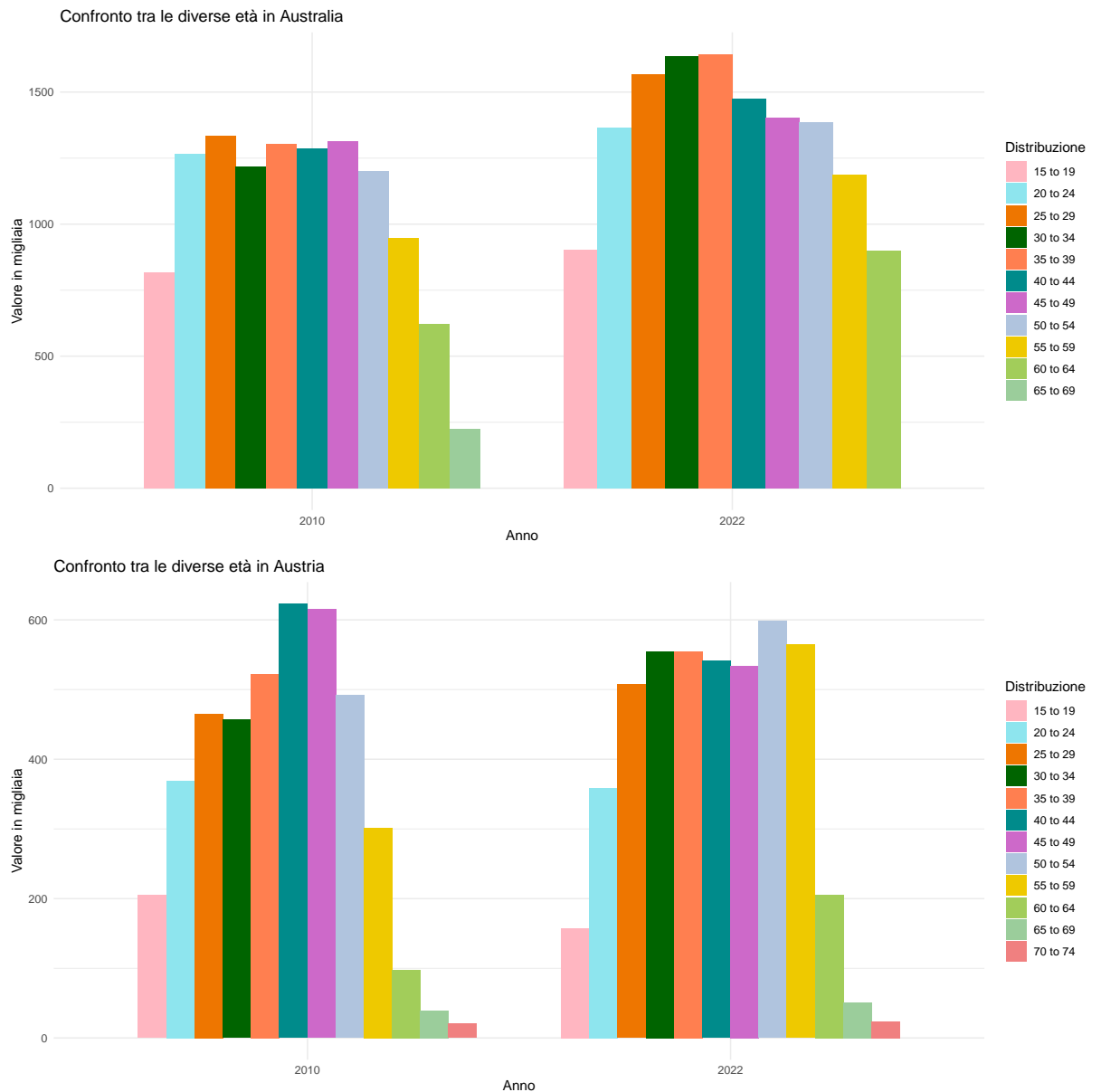
Inoltre, la *gender inequality* persiste in molti settori del mercato del lavoro, influenzando le condizioni di lavoro, le possibilità di promozione e le retribuzioni [16]. Le donne possono essere più inclini ad accettare lavori precari o part-time a causa della limitata accessibilità a lavori a tempo pieno ben retribuiti. Questo può influenzare la differenza nei tassi di disoccupazione tra uomini e donne. Infatti, le responsabilità familiari e di cura possono limitare la disponibilità delle donne a lavorare a tempo pieno o in posizioni che richiedono orari flessibili.

Nel grafico precedente, dal 2010 al 2022, si può notare come la discrepanza tra i disoccupati maschi e i disoccupati femmine si sta gradualmente appianando. Ciò potrebbe significare che anche gli uomini stanno iniziando a fare sempre più lavori che normalmente sarebbero destinati alle donne, e quindi la differenza di disoccupazione sta diminuendo.

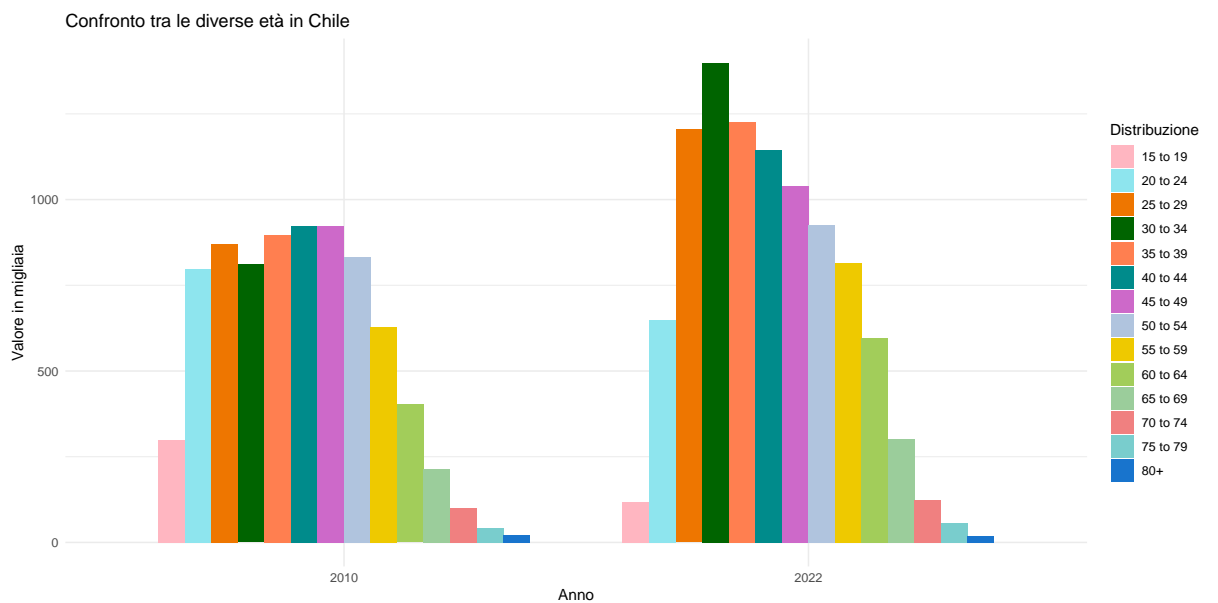
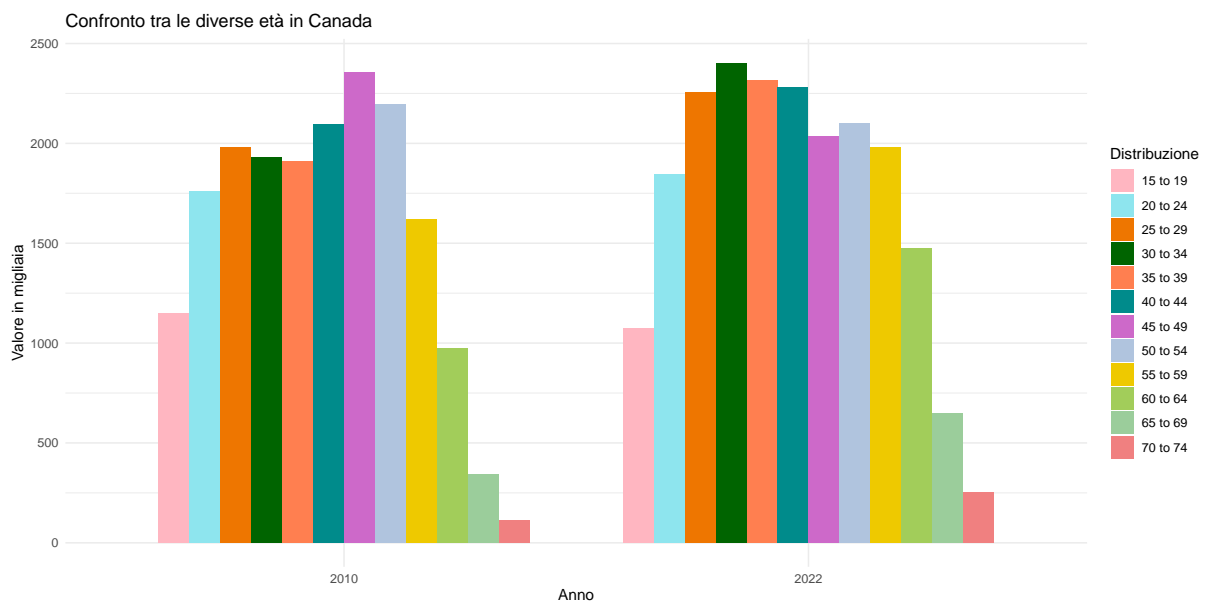
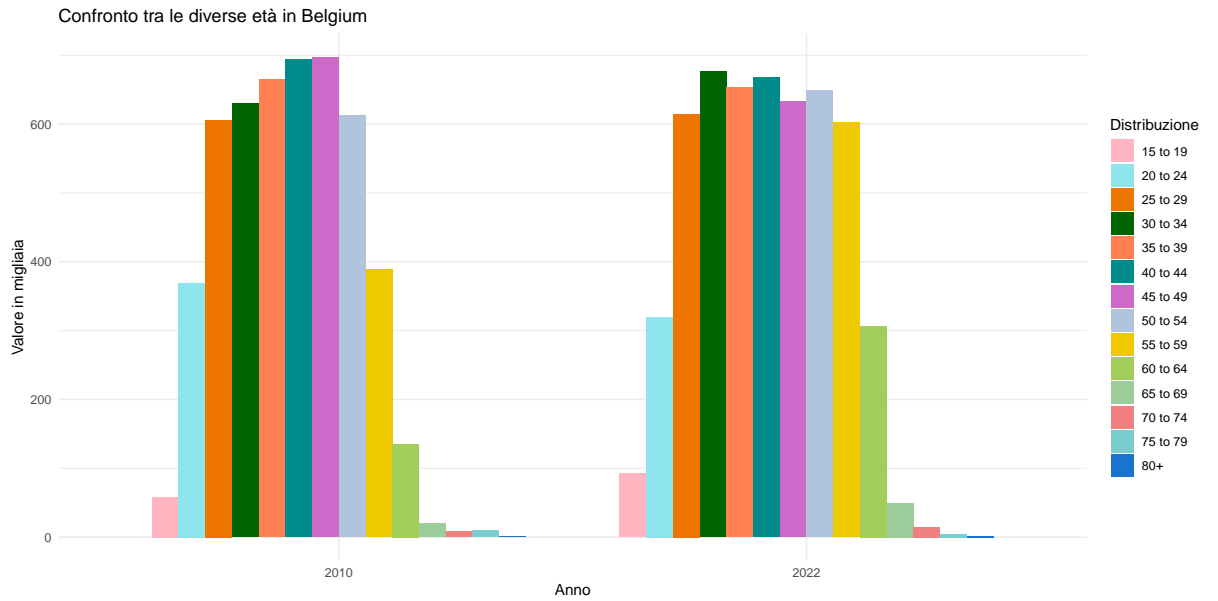
In conclusione, la combinazione di questi fattori può contribuire a una maggiore propensione delle donne a occupare lavori part-time o mal pagati, che possono offrire maggiore flessibilità ma anche minore sicurezza e retribuzione. Questa situazione può influenzare i tassi di disoccupazione, con meno donne che risultano ufficialmente disoccupate rispetto agli uomini, ma con una presenza più diffusa in lavori con condizioni di lavoro precarie.

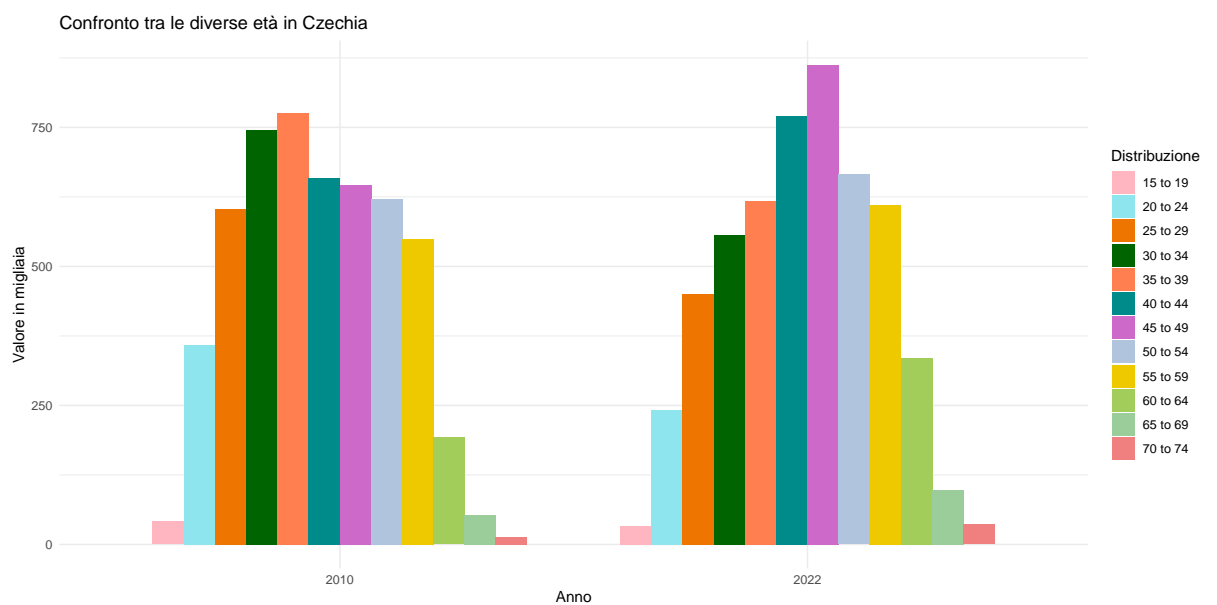
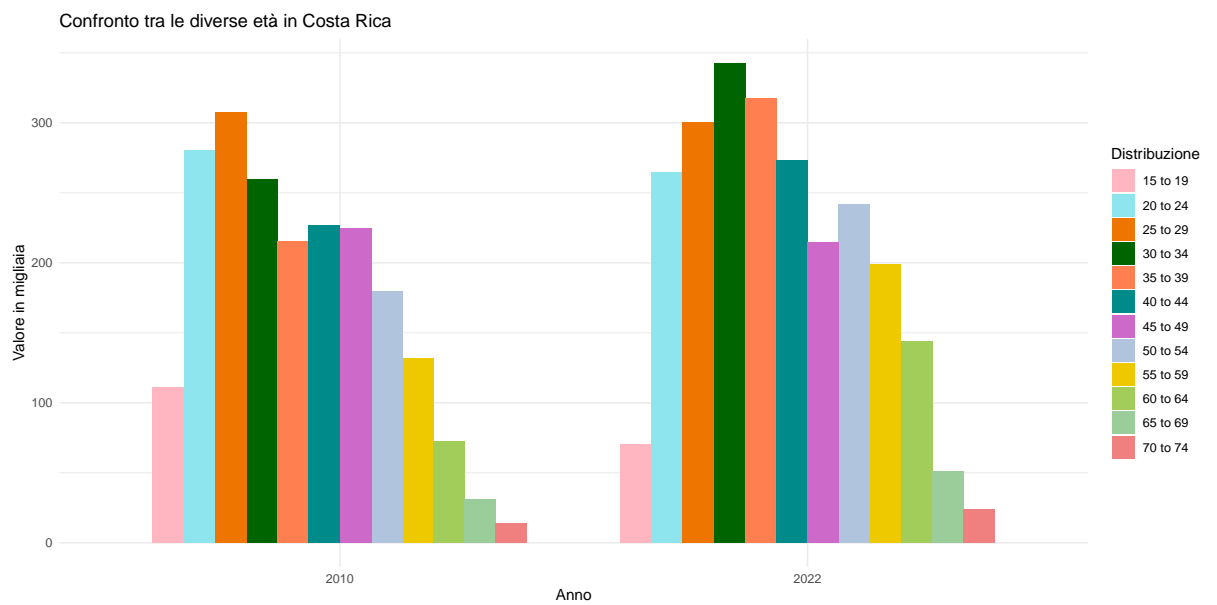
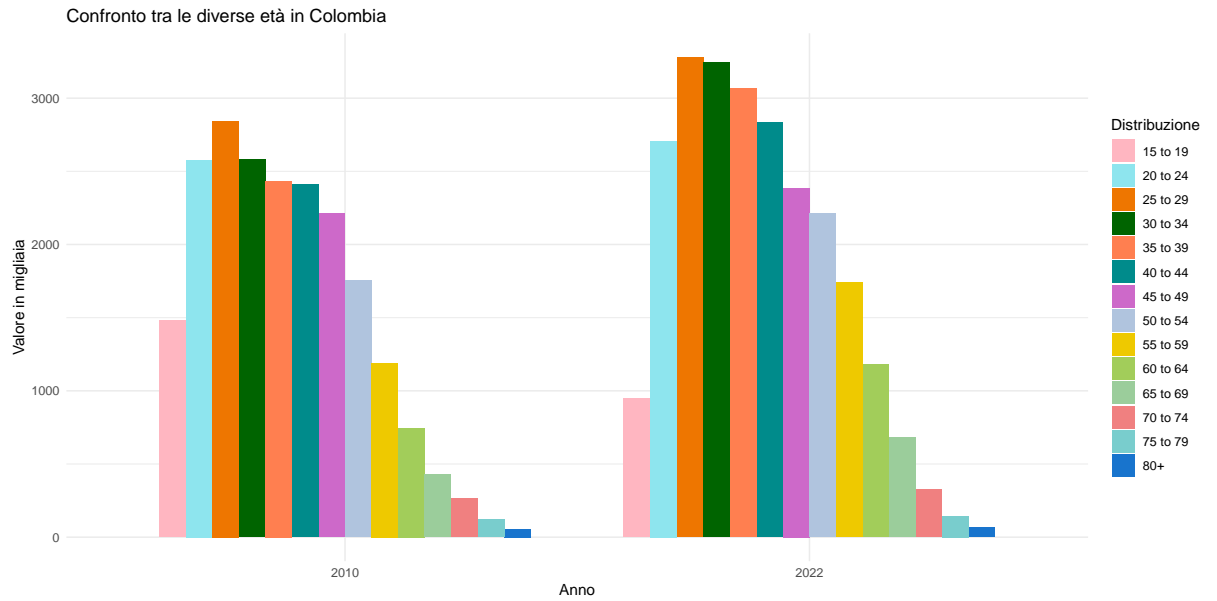
## 3.2 Barplot sull'età

Di seguito sono riportati i barplot (di ogni paese) per categoria d'età, nel 2010 e nel 2022, per fare un confronto tra le due distribuzioni. A destra del grafico c'è la legenda con tutte le categorie di età che vengono considerate.

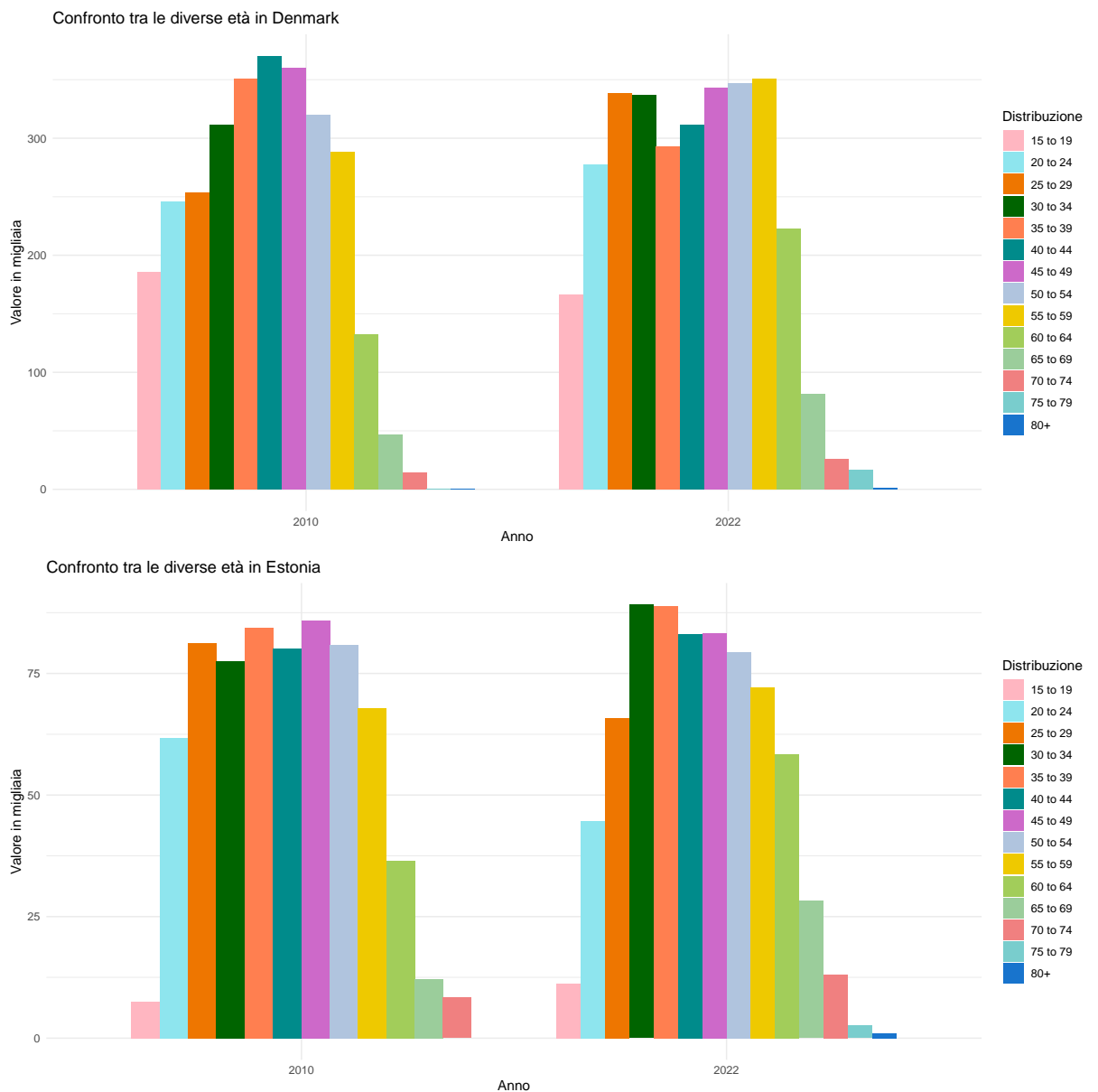


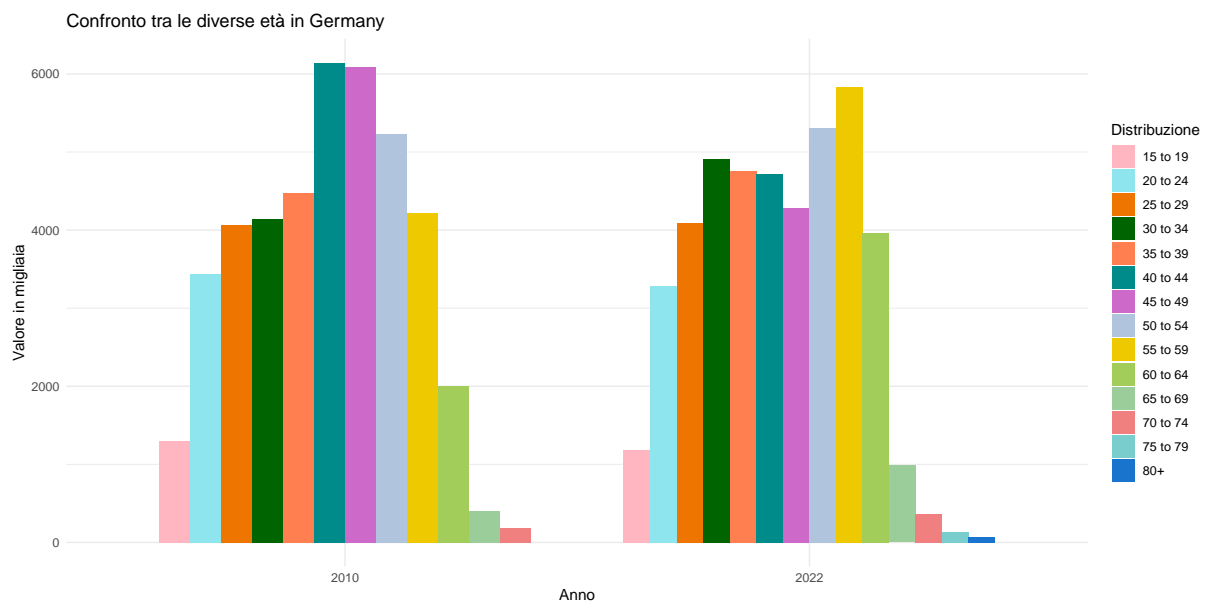
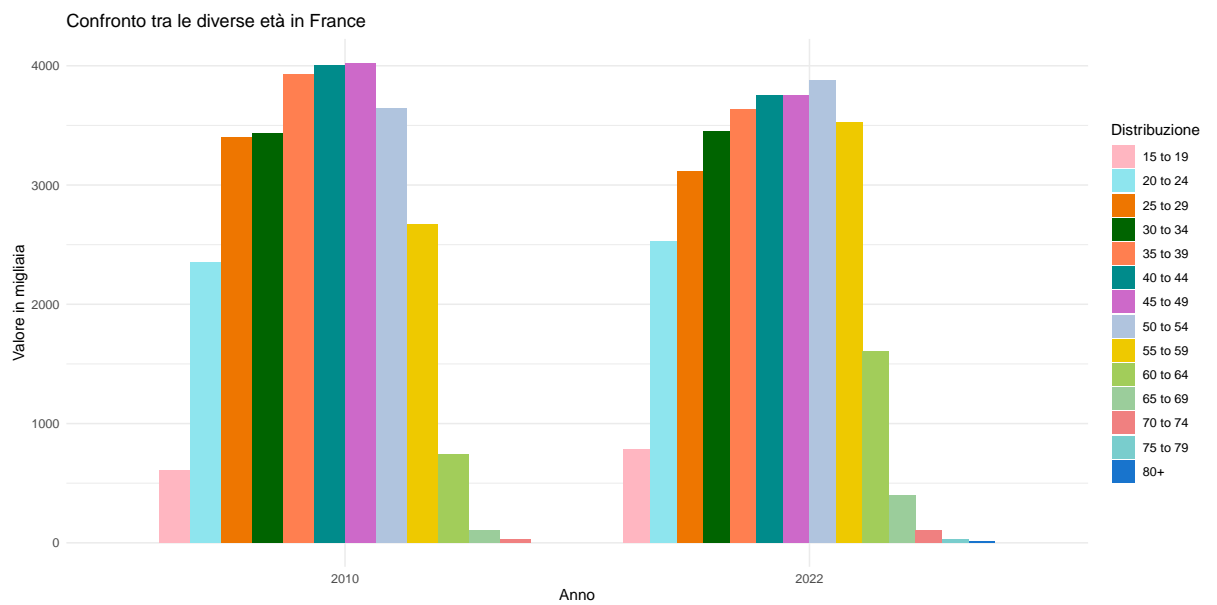
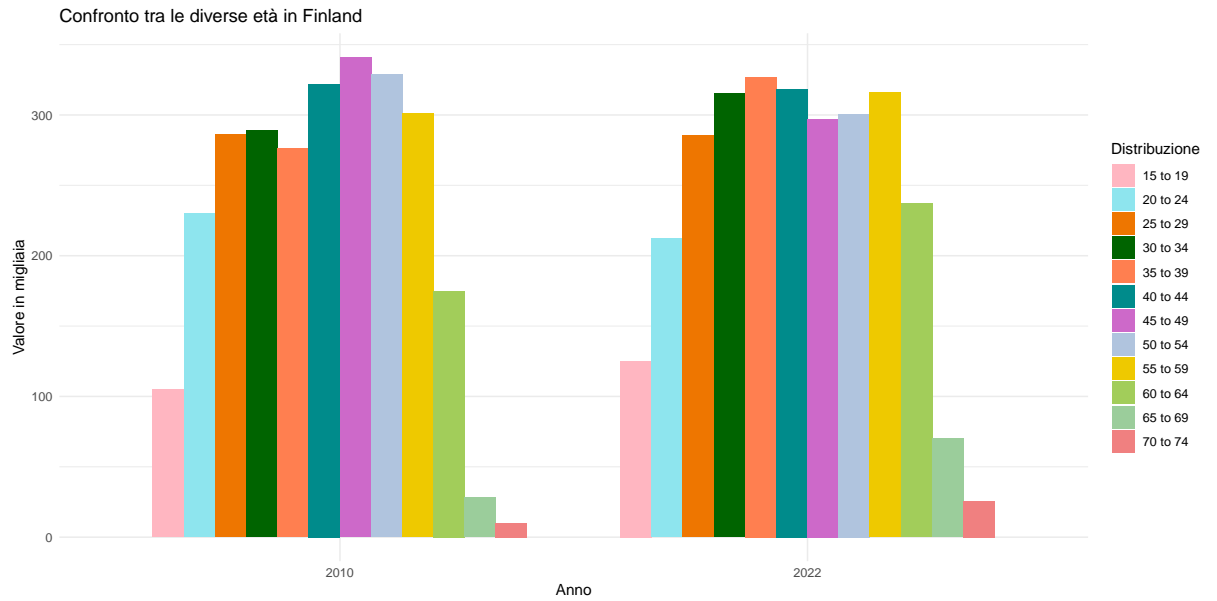




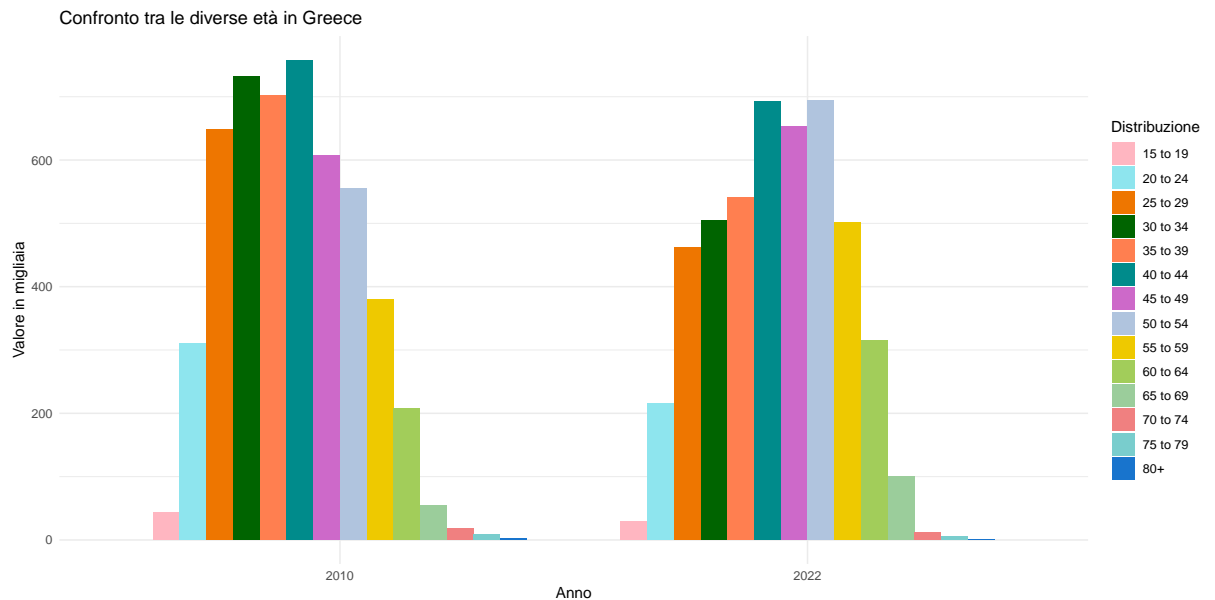


In Repubblica Ceca si osserva un'invecchiamento della popolazione, in parte a causa di un tasso di natalità relativamente basso [17] e di una crescente aspettativa di vita. Inoltre, la Repubblica Ceca ha subito un periodo di emigrazione dopo la caduta del comunismo [18], il che ha contribuito ad un calo della popolazione giovanile.

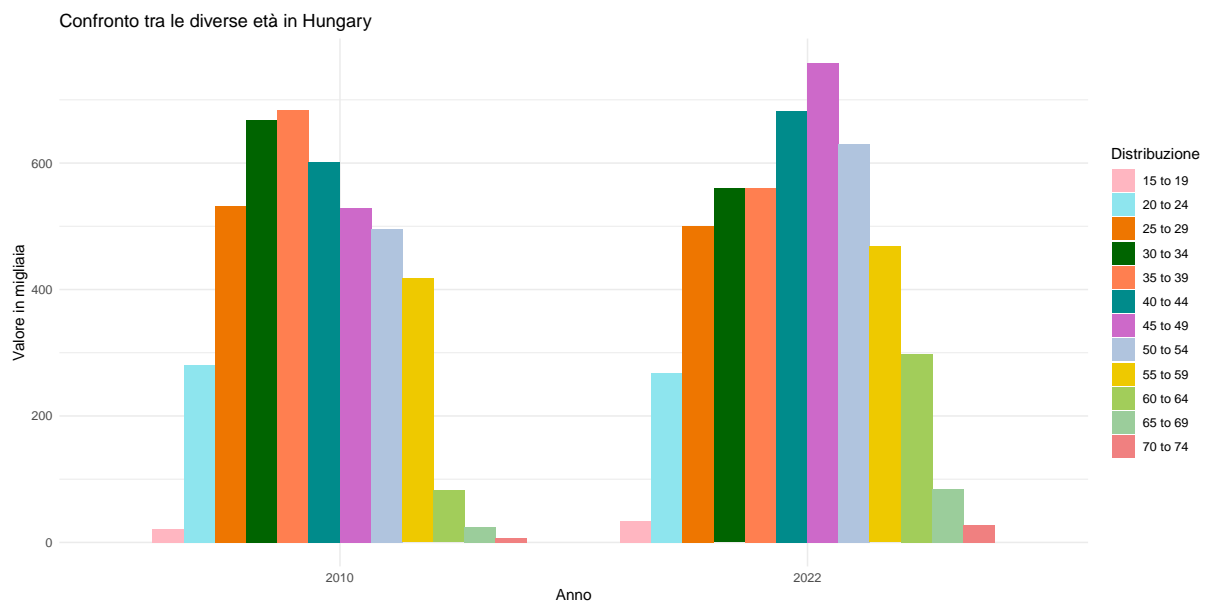


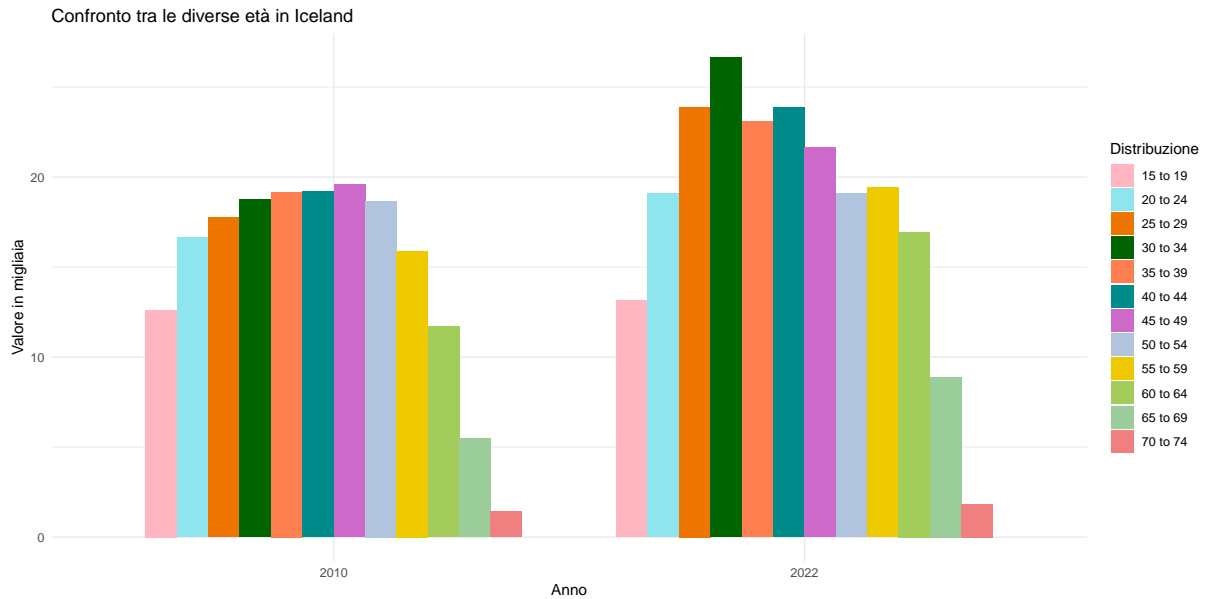


La Germania ha una delle popolazioni più invecchiate al mondo, con una bassa natalità e una popolazione che sta vivendo più a lungo grazie ai progressi nella medicina e nella salute pubblica [19].

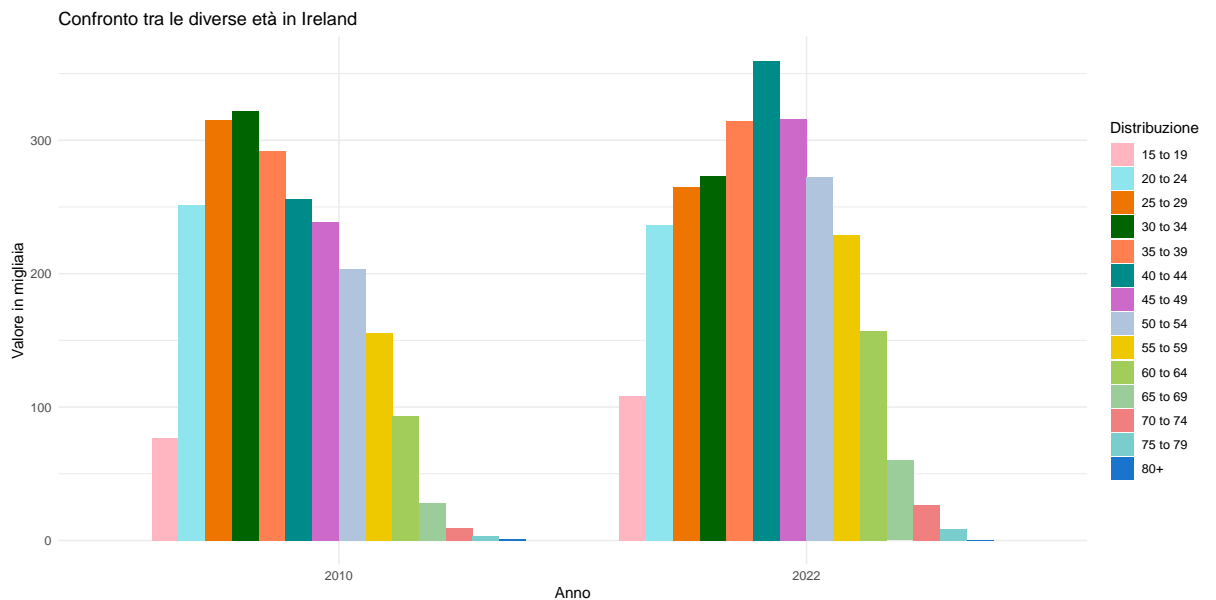


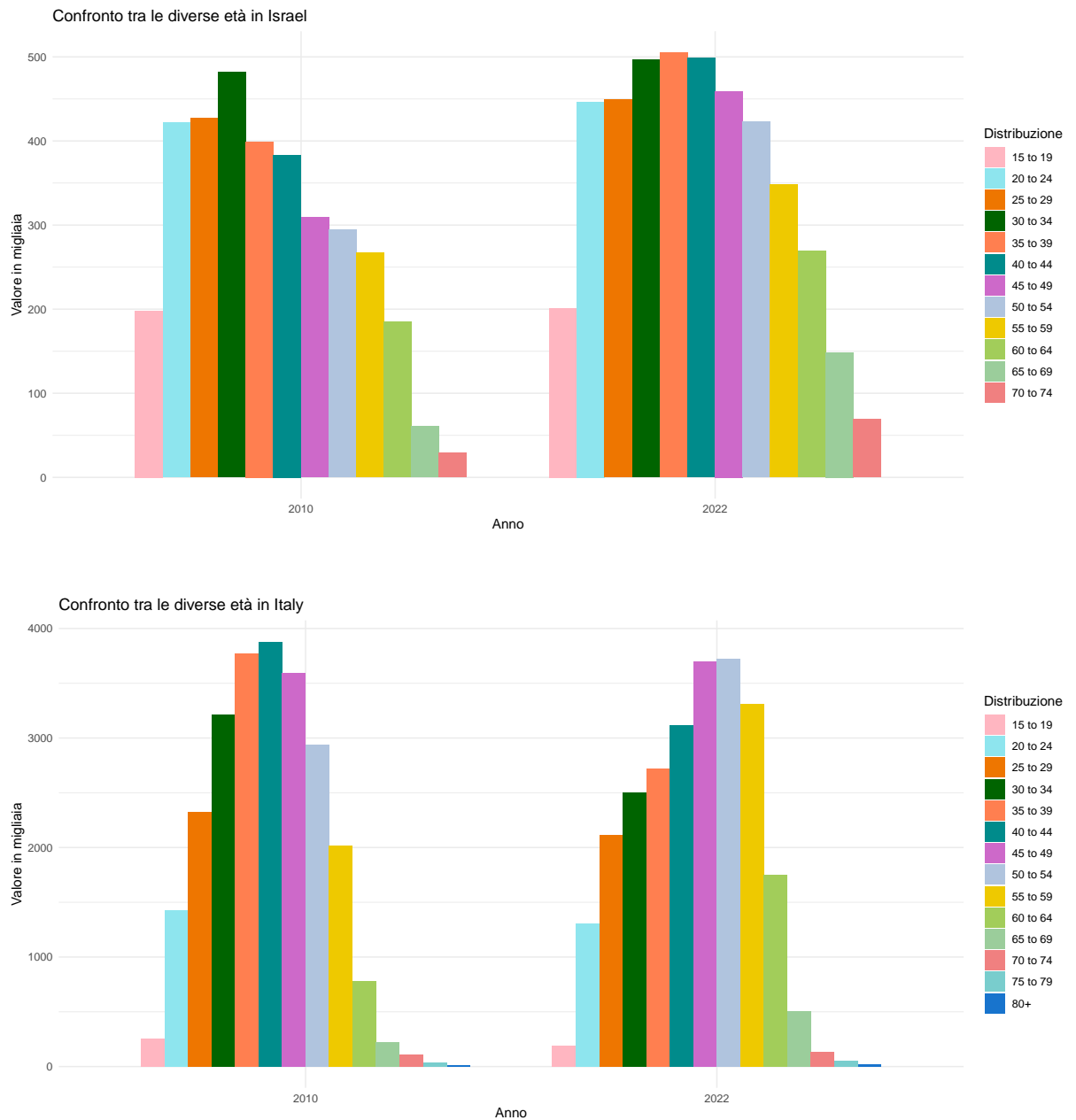
La Grecia ha un tasso di natalità in calo e una popolazione che sta vivendo più a lungo grazie ai progressi nella sanità e nella medicina. Inoltre, la crisi economica che ha colpito la Grecia negli ultimi anni ha influenzato anche il tasso di natalità e ha portato a una diminuzione della popolazione giovanile [20].



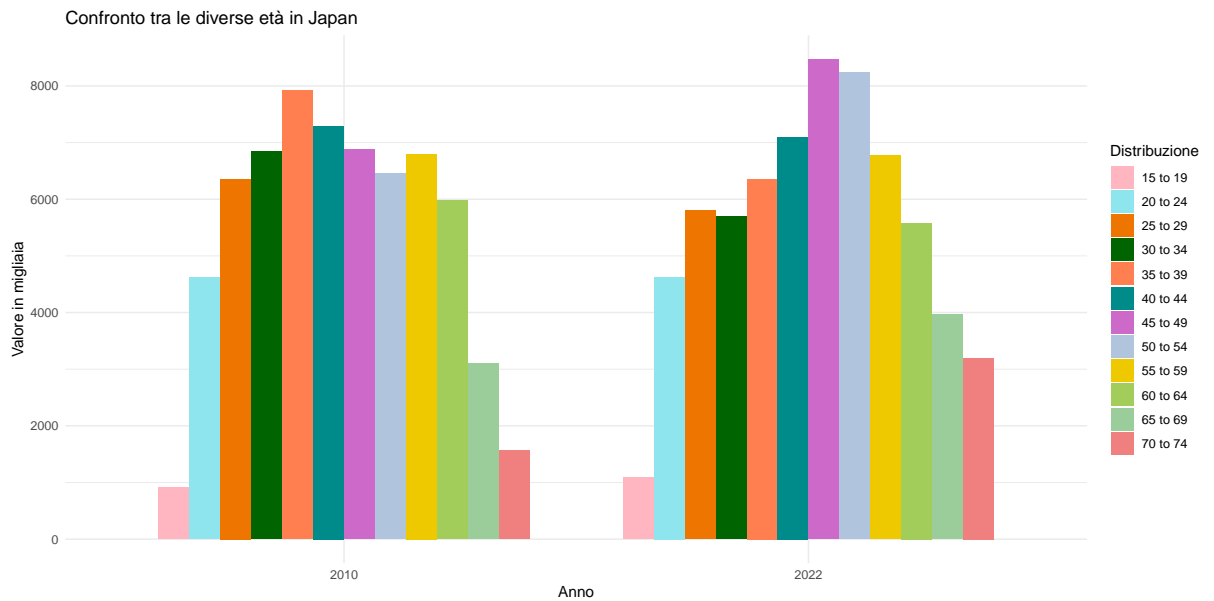


L'Islanda registra un aumento della popolazione giovane. I motivi sono: Politiche governative che sostengono le famiglie e incentivano la natalità; Una buona qualità della vita, che incoraggia le persone a stabilirsi in questi paesi e a formare famiglie. Per queste ragioni bisogna considerare anche l'aspetto dell'immigrazione: l'Islanda potrebbe attrarre immigrati giovani e lavoratori qualificati da altre parti del mondo [21].

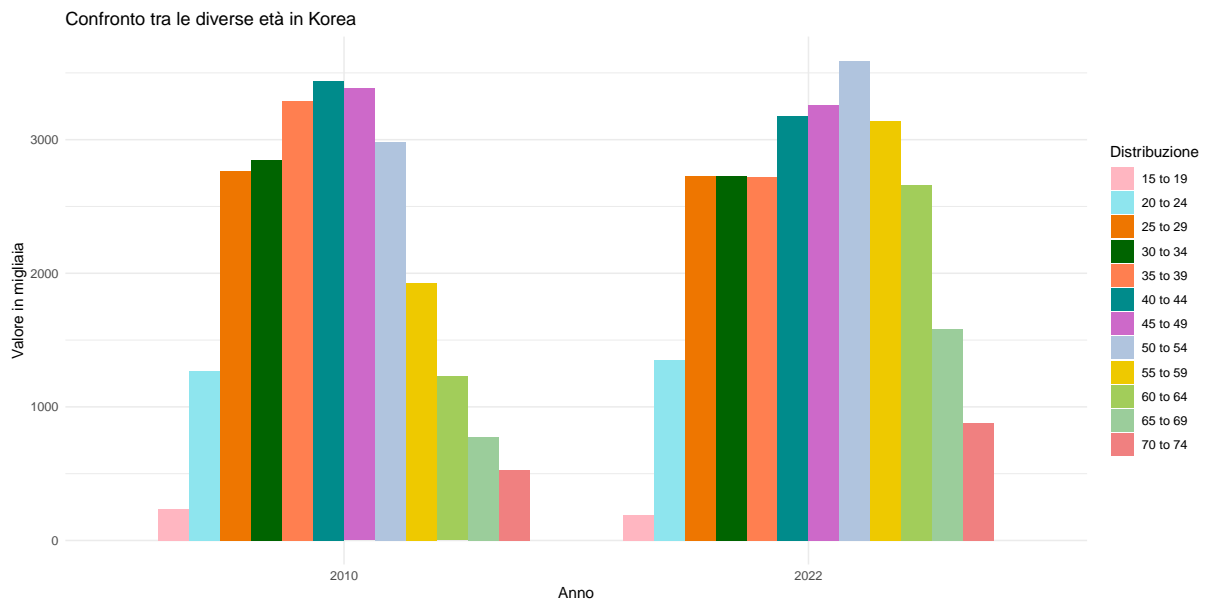




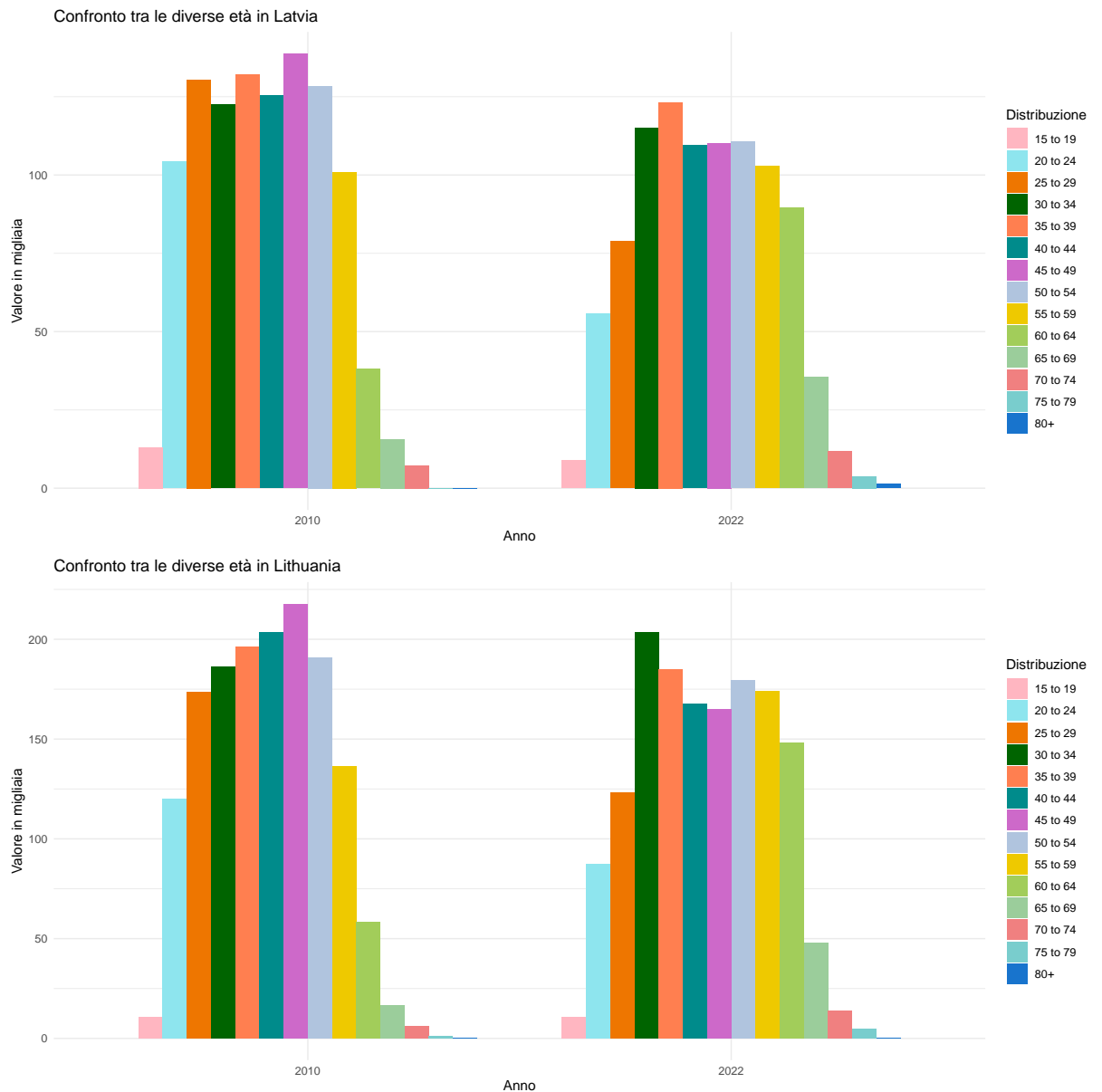
L'italia è il paese che presenta l'invecchiamento della popolazione più alto di tutti i paesi. Ciò è dovuto a una bassa natalità, un aumento dell'aspettativa di vita e all'emigrazione di giovani all'estero in cerca di lavoro. In Italia, il 2022 si contraddistingue per un nuovo record del minimo di nascite (393mila, per la prima volta dall'Unità d'Italia sotto le 400mila) e per l'elevato numero di decessi (713mila) [22].



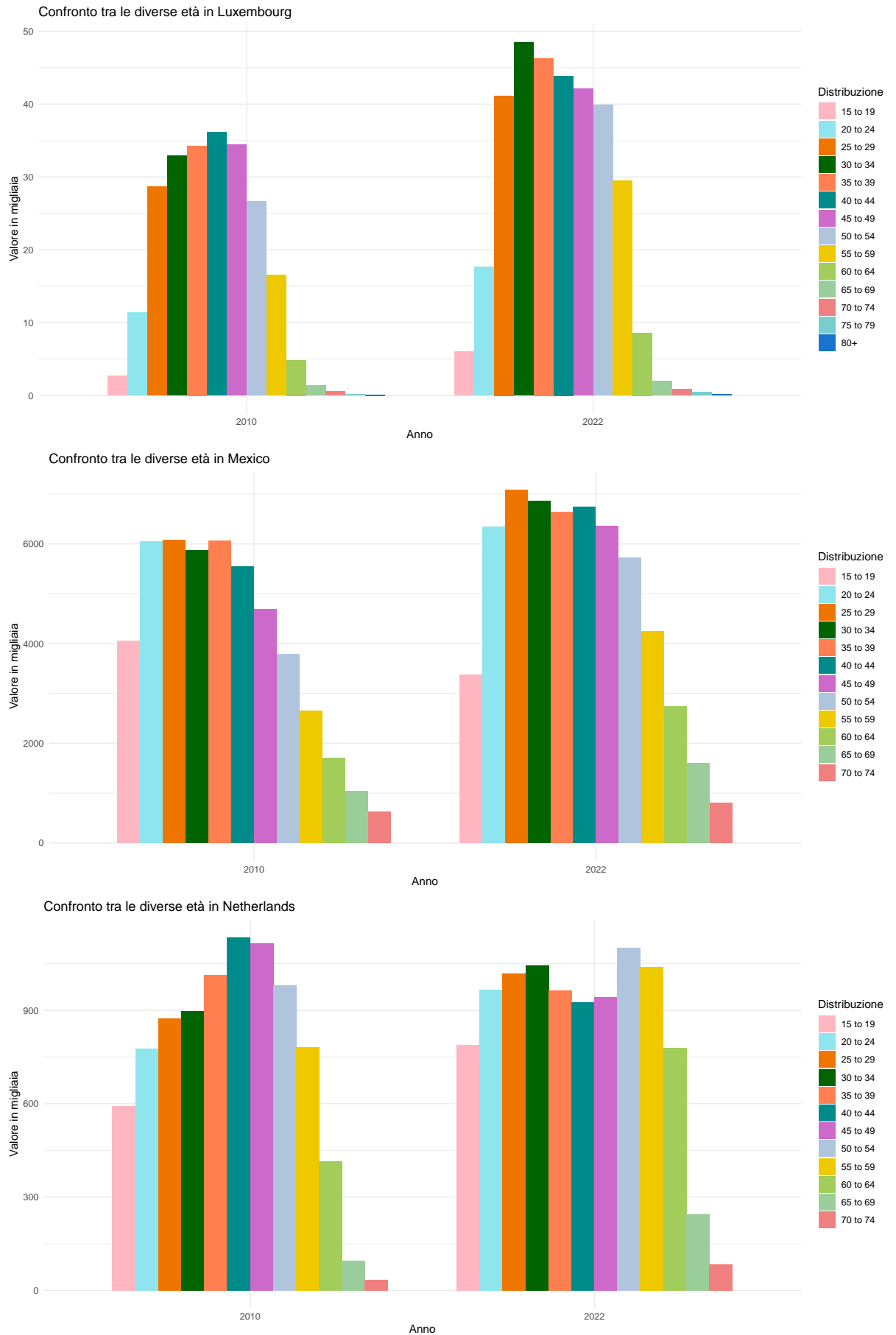
Il Giappone ha uno dei tassi di natalità più bassi al mondo. Questo è attribuibile a diversi fattori: l'alto costo della vita [23], la difficoltà di conciliare il lavoro con la famiglia [24], la mancanza di politiche di sostegno alla famiglia [25]. Tutto ciò va in piena contrapposizione con la politica del modello svedese (che sarà analizzata approfonditamente nella sezione "Cluster K-MEANS").



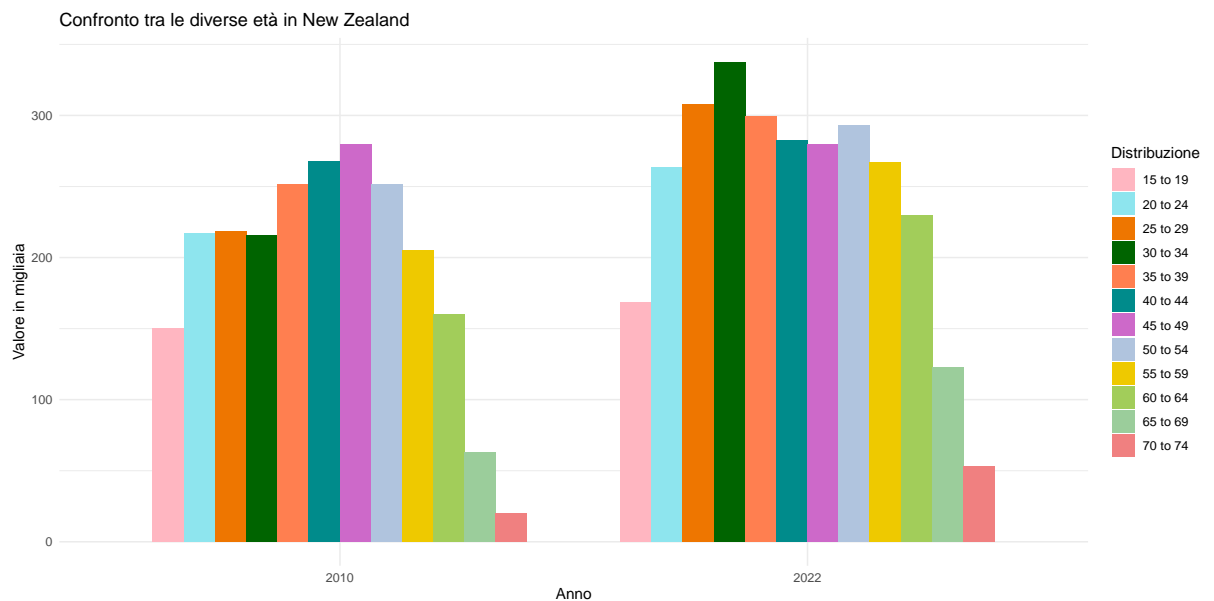




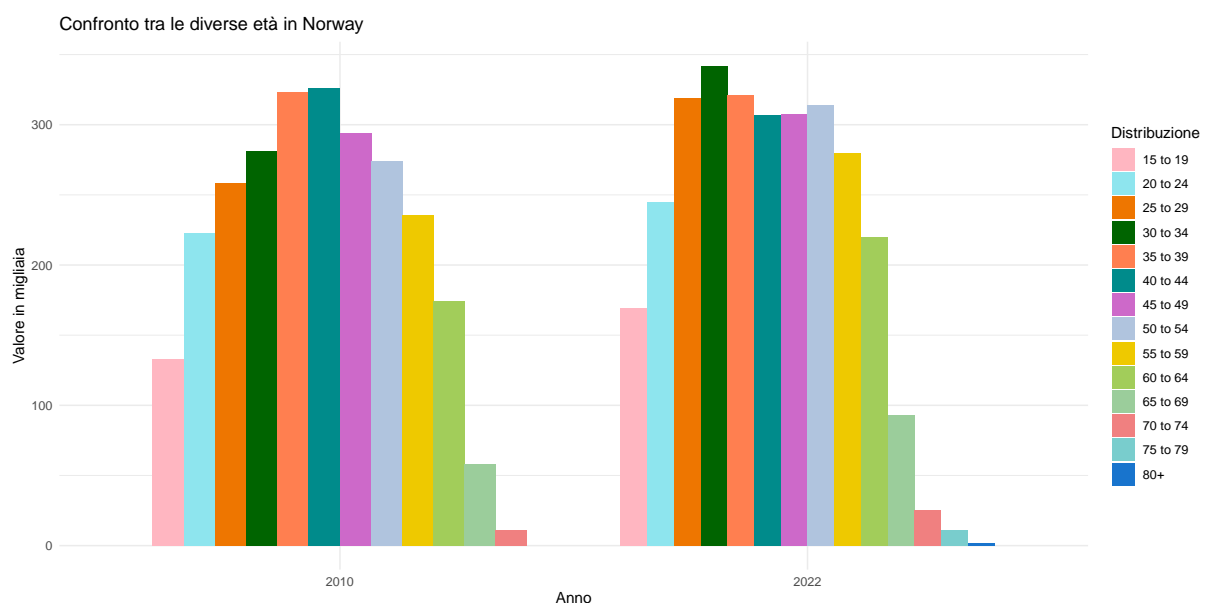
La Lituania presenta sia un invecchiamento che un ringiovanimento della popolazione. L'invecchiamento è dovuto all'aumento dell'aspettativa di vita, mentre il ringiovanimento è causato da un "baby boom" nel lontano 2009 [26], anche se con il passare degli anni il tasso di fertilità è diminuito drasticamente fino ad arrivare a 1.36 nel 2022. Quindi in realtà la Lituania ha registrato un aumento delle nascite grazie agli anni precedenti, ma deve risolvere al più presto il problema della bassa fertilità per evitare un eccessivo invecchiamento della popolazione.

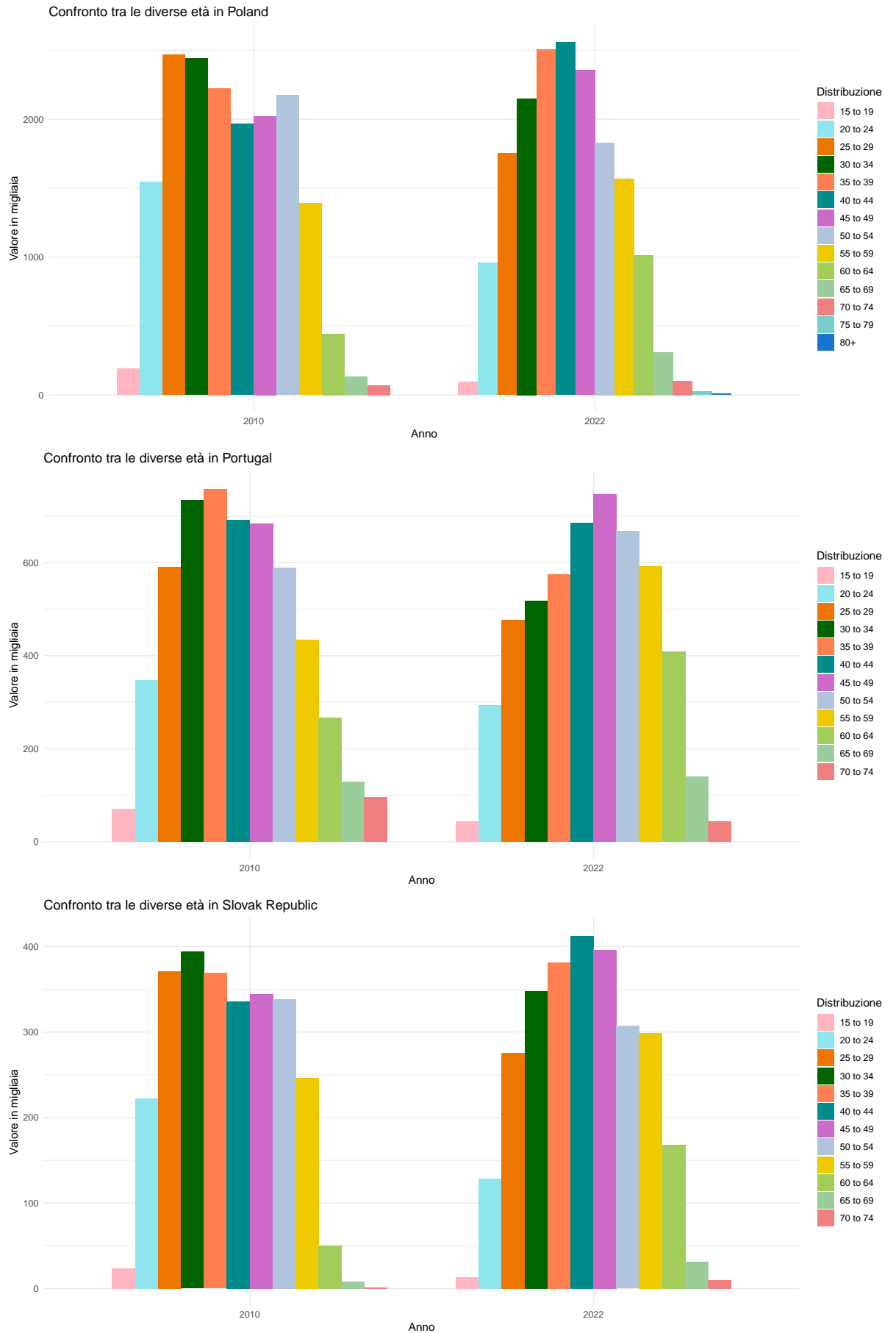


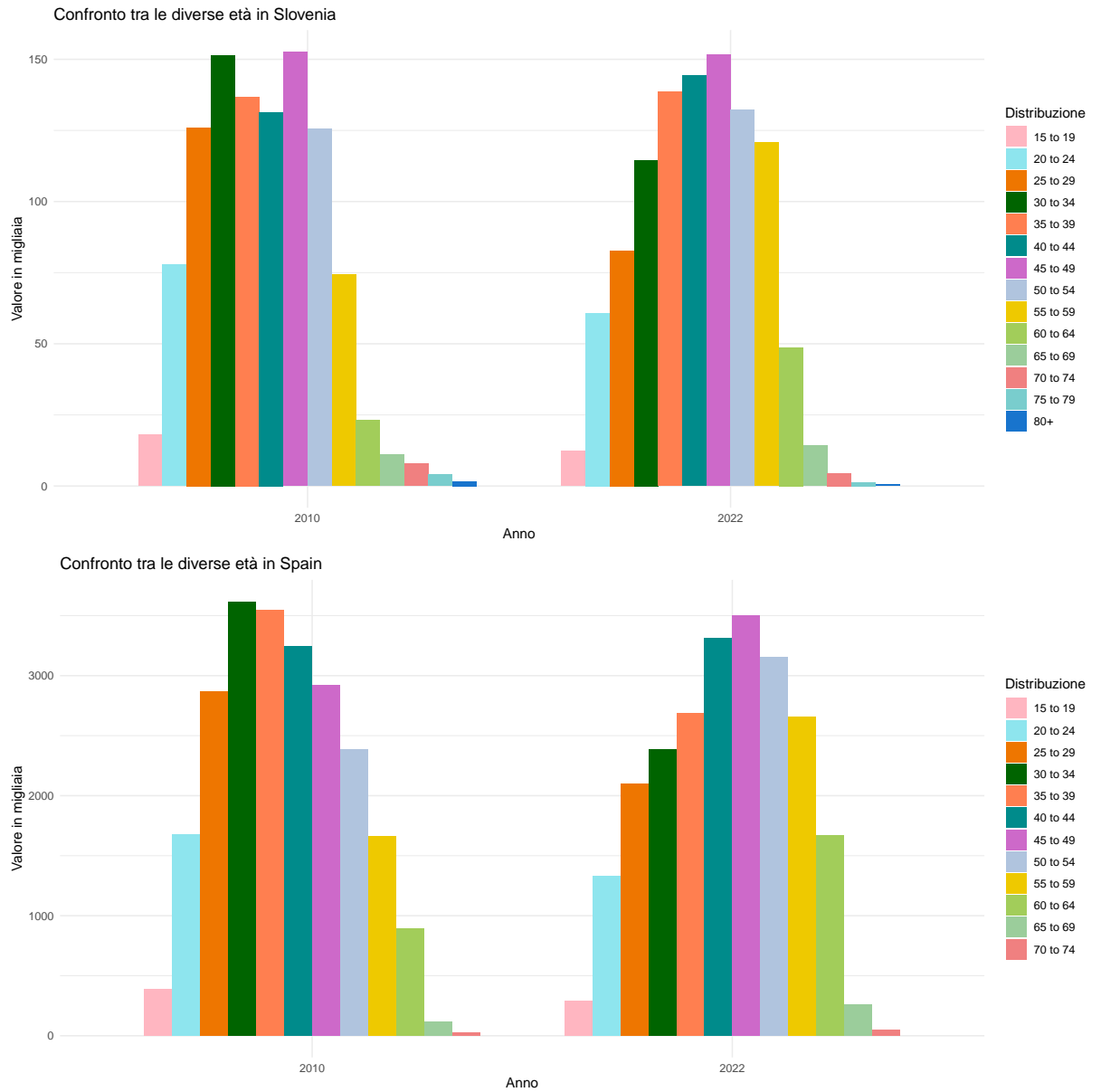
L'Olanda, come la Lituania, presenta sia un invecchiamento che un ringiovanimento della popolazione. L'invecchiamento è comune alla maggior parte dei paesi ed è dovuto ad un aumento dell'aspettativa di vita. Il ringiovanimento è probabilmente dovuto a politiche di sostegno per la cura dei figli, e all'immigrazione nel paese di lavoratori giovani e qualificati.



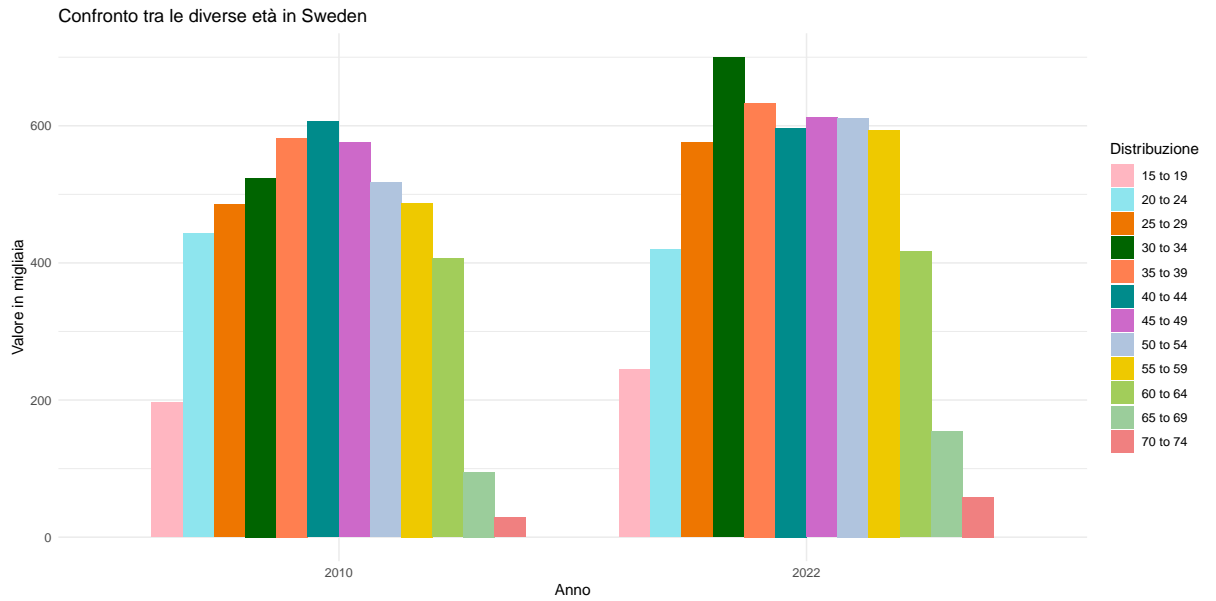
La Nuova Zelanda presenta lo stesso ringiovanimento della popolazione dell'Islanda, per motivi altrettanto simili.



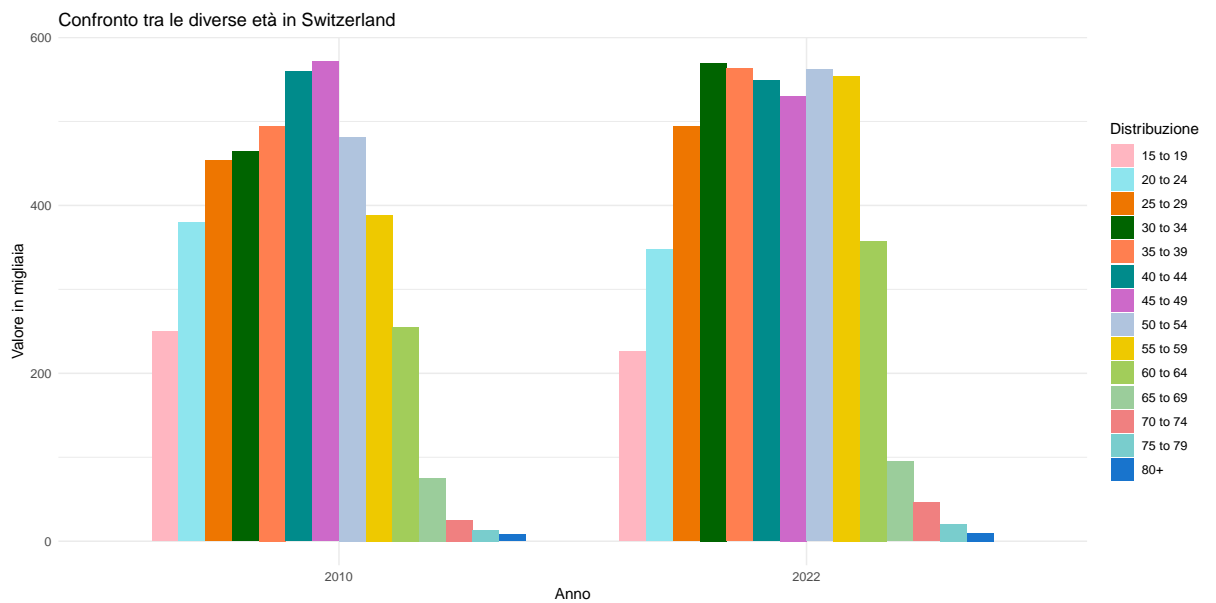


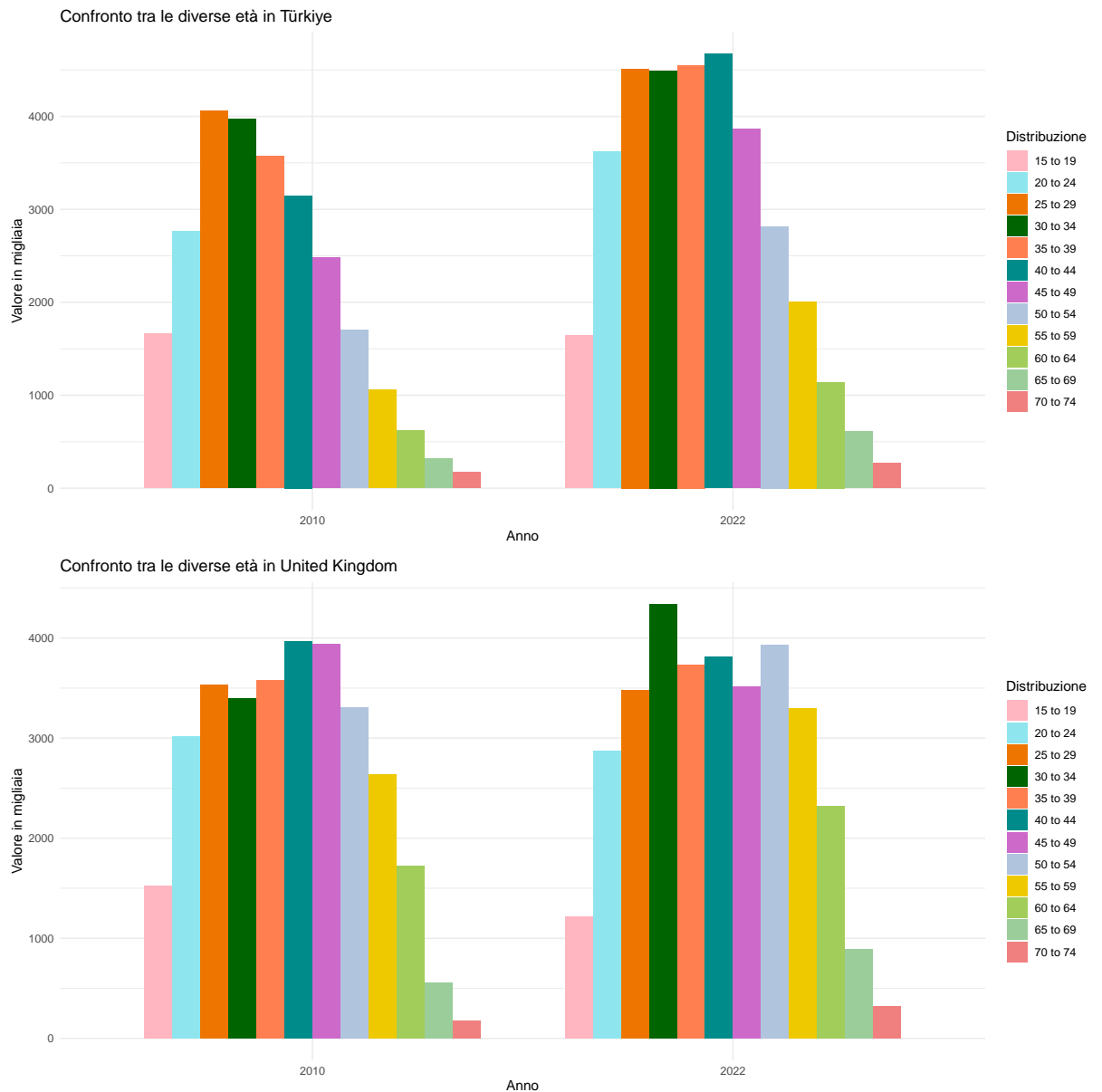


La Spagna presenta le stesse variazioni e gli stessi problemi della Grecia, dovuti alla crisi finanziaria del 2008.

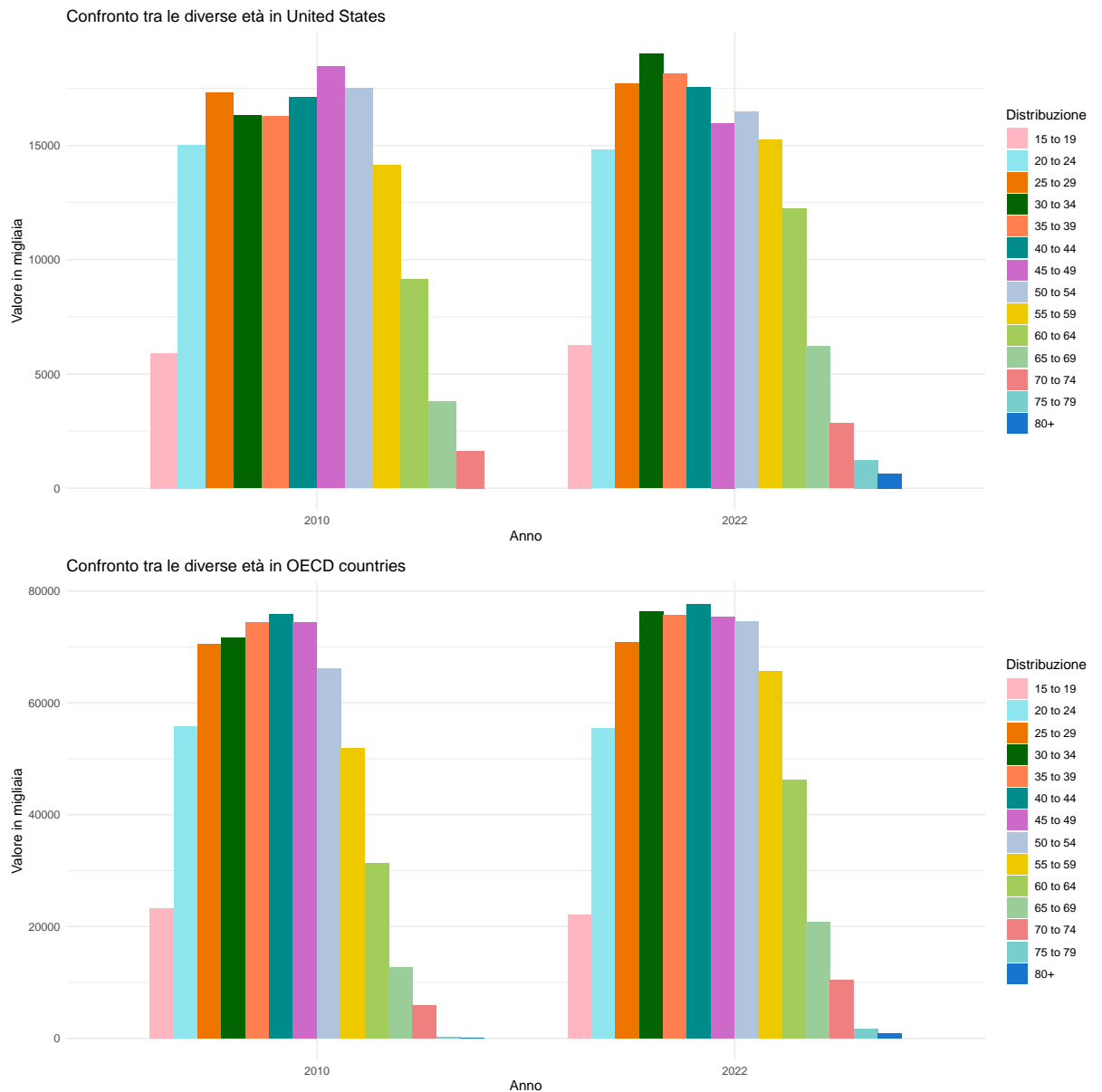


La Svezia presenta le stesse variazioni dell'Islanda. Questi due paesi sono infatti quelli che più hanno beneficiato nell'adottare la politica del modello svedese.





Il Regno Unito mostra sia un invecchiamento che un ringiovanimento della popolazione, come l'Olanda e la Lituania. I motivi sono probabilmente gli stessi dell'Olanda, a causa delle politiche adottate molto simili.



Per quasi tutti i paesi si può notare un incremento generale della popolazione. Ciò è dovuto al fatto che nel 2010 la popolazione mondiale era costituita da 7 miliardi di persone, mentre nel 2022 è arrivata fino a 8 miliardi. Inoltre, l'incremento di persone anziane è dovuto all'aumento dell'aspettativa di vita [27].



## CAPITOLO 4

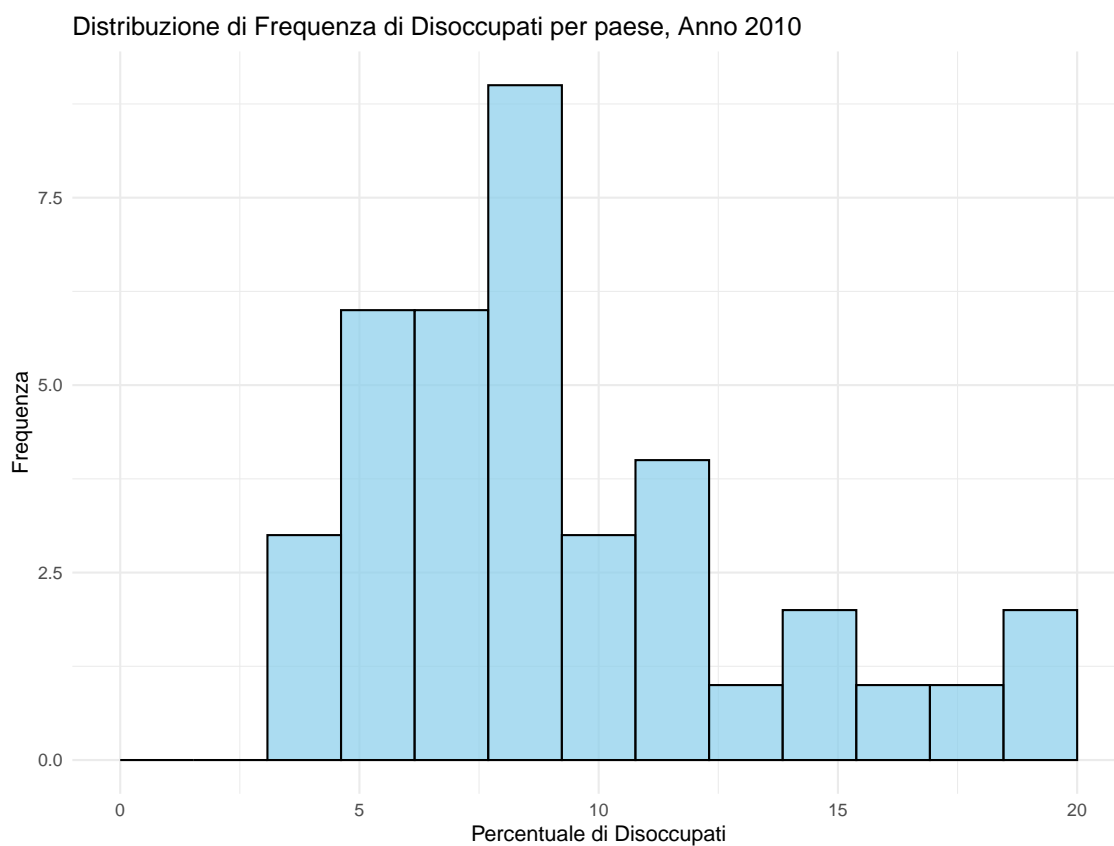
---

### Istogrammi e Boxplot

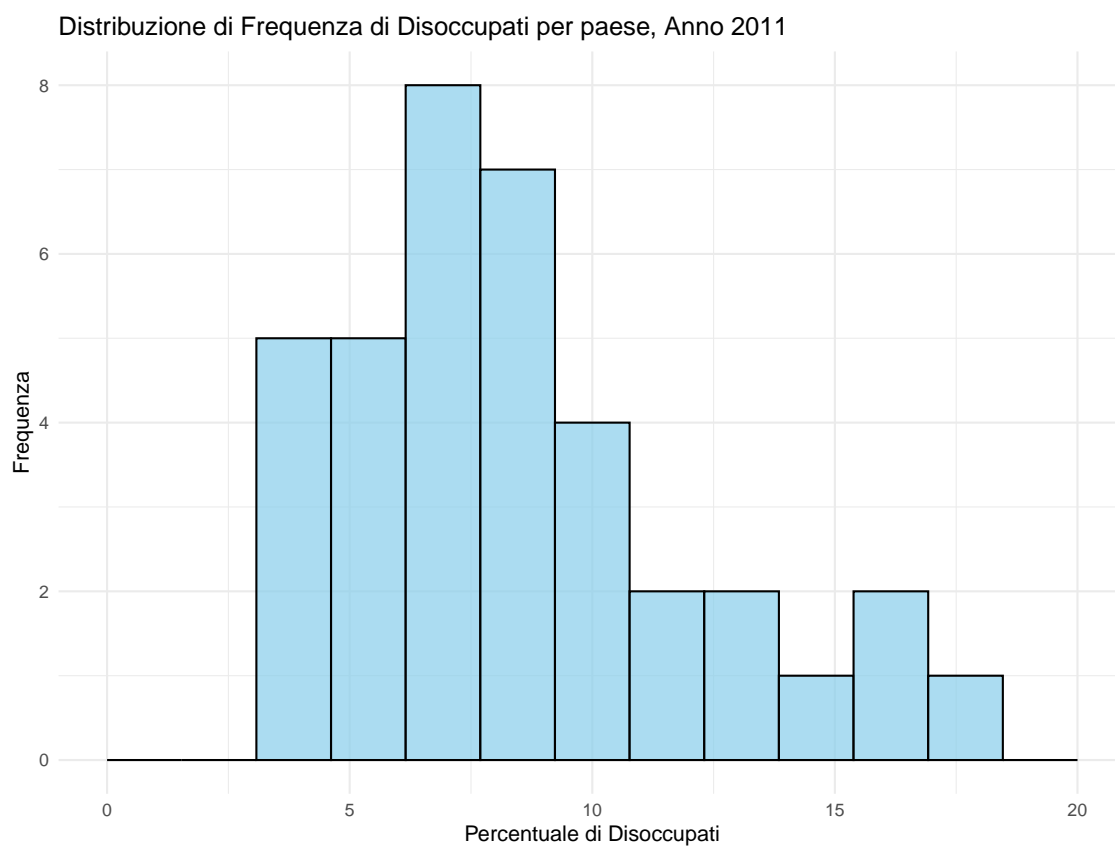
---

#### 4.1 Istogrammi

Di seguito sono riportati gli istogrammi, dall'anno 2010 al 2022, delle frequenze delle percentuali di disoccupazione (rispetto alla forza lavoro) dei 38 paesi. Per migliorare la leggibilità, per ogni grafico appaiono un massimo di 13 barre, e l'intervallo delle ascisse va dallo 0% al 20% o 28%, a seconda degli anni. Più precisamente, dal 2013 al 2017, i valori delle ascisse partono da 0 fino ad arrivare a 28, per mostrare la presenza di outlier. Tutti gli altri anni i valori delle ascisse vanno da 0 a 20.

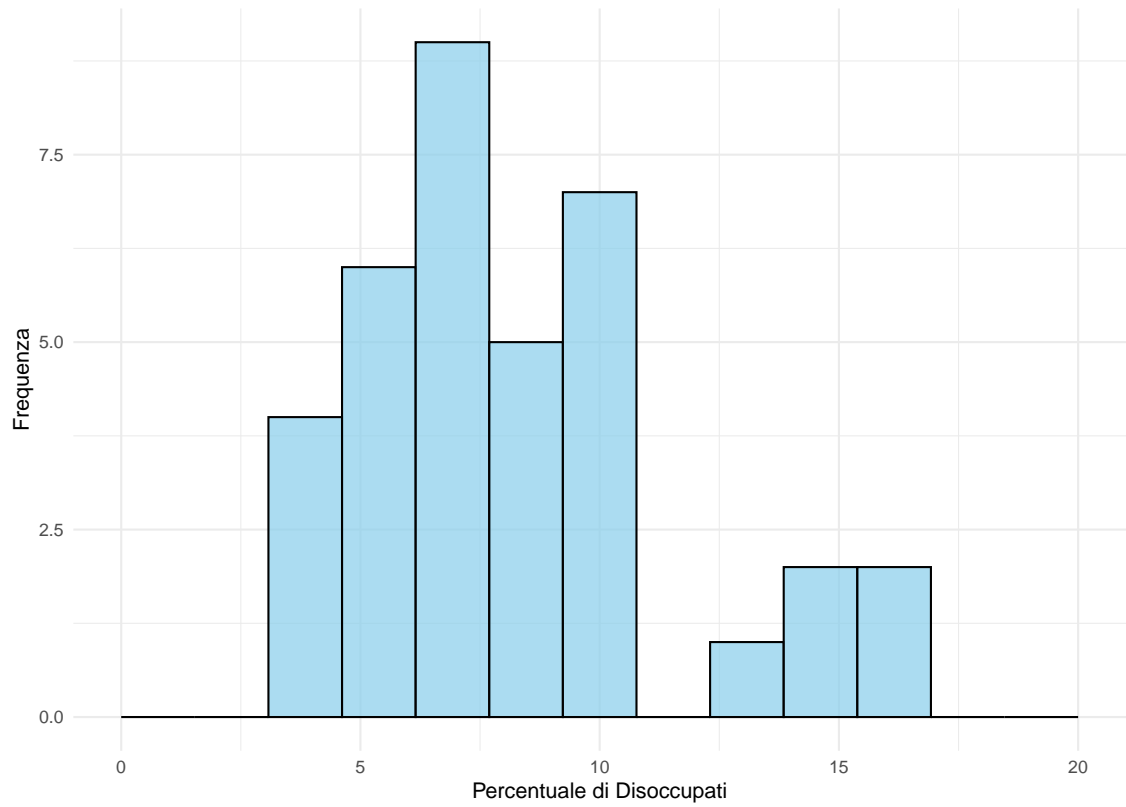


Nel 2010 si può osservare un picco intorno al 9%, che riguarda 9 paesi. Inoltre, ben 6 paesi hanno una percentuale di disoccupazione che va dal 15% al 20%. Probabilmente questo è attribuibile alla crisi finanziaria del 2008.

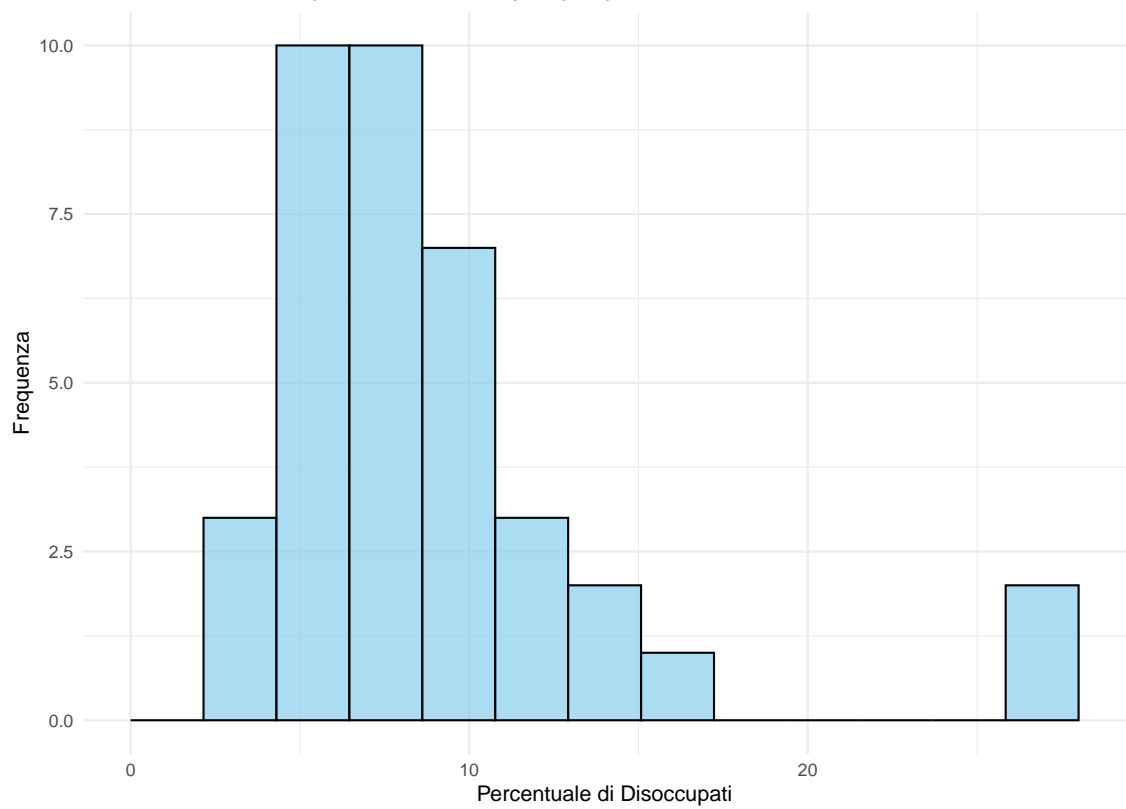


Nel 2011 il picco è intorno al 7%, che vien raggiunto da 8 paesi. Nessun paese ha più una percentuale di disoccupazione del 20%.

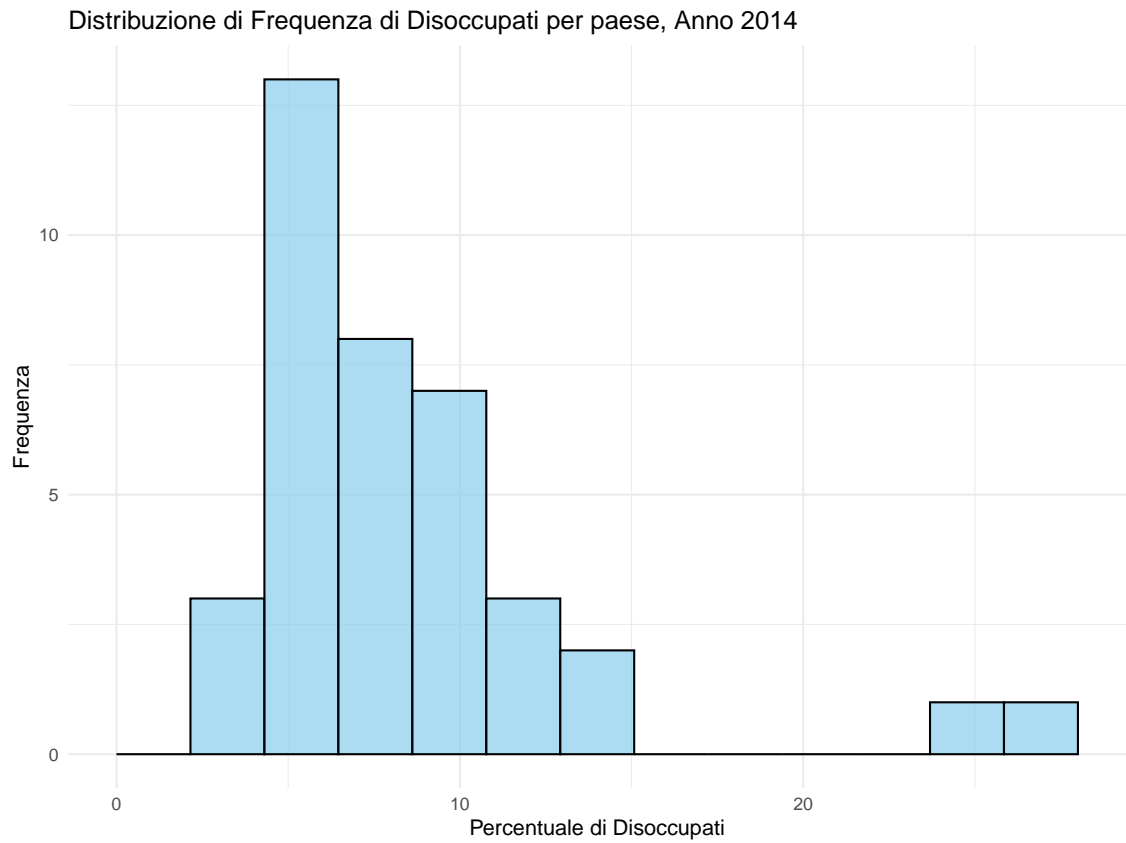
Distribuzione di Frequenza di Disoccupati per paese, Anno 2012

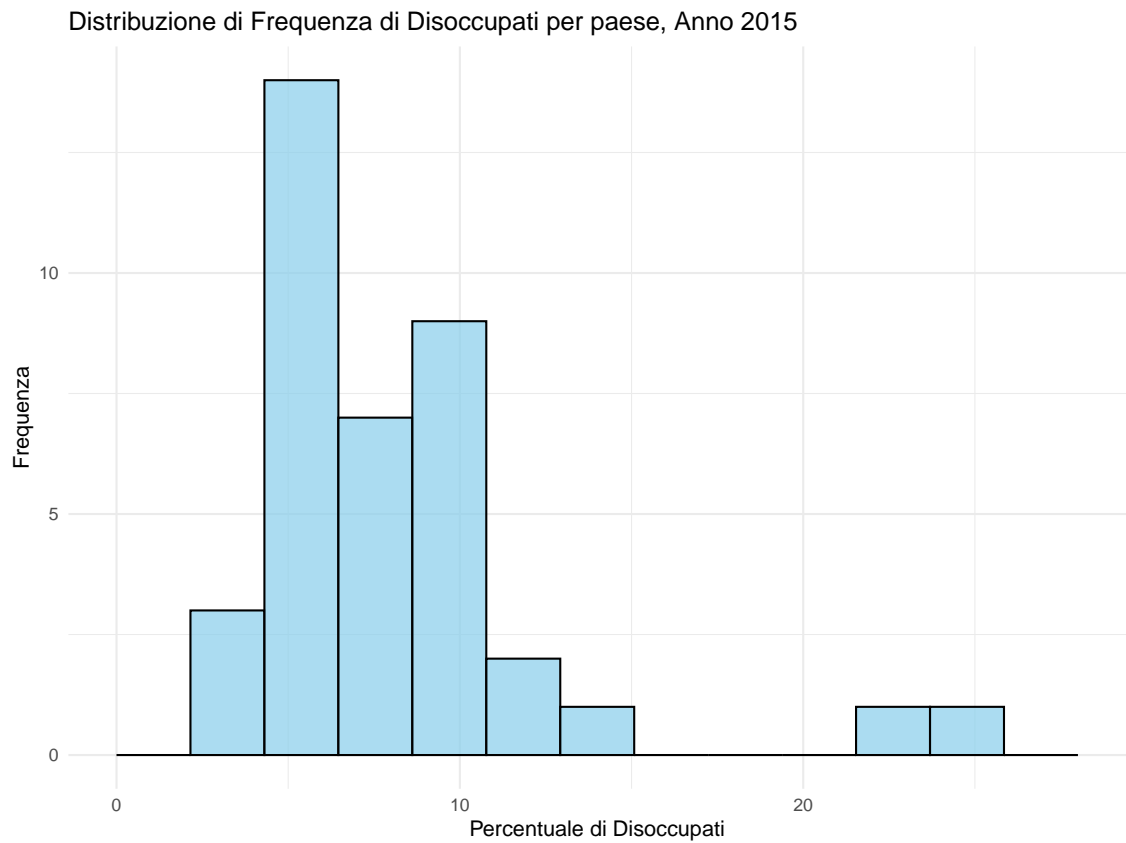


Distribuzione di Frequenza di Disoccupati per paese, Anno 2013

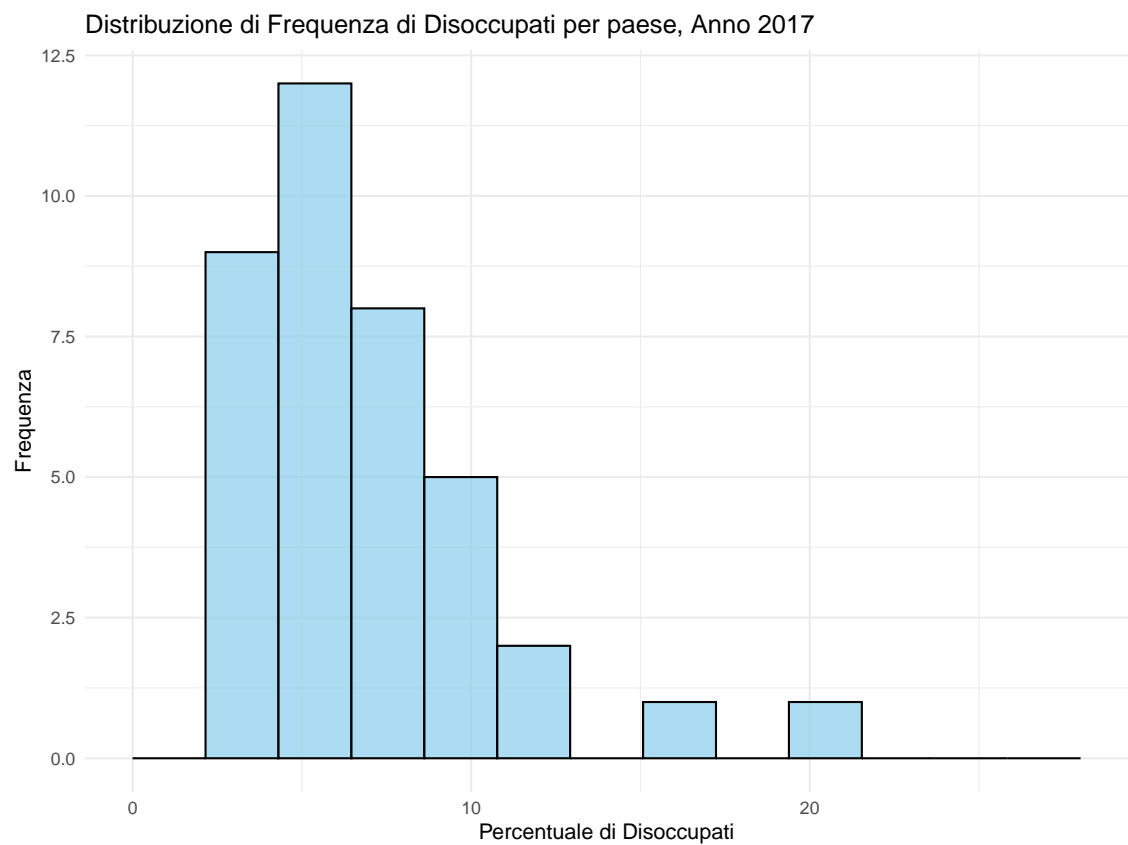
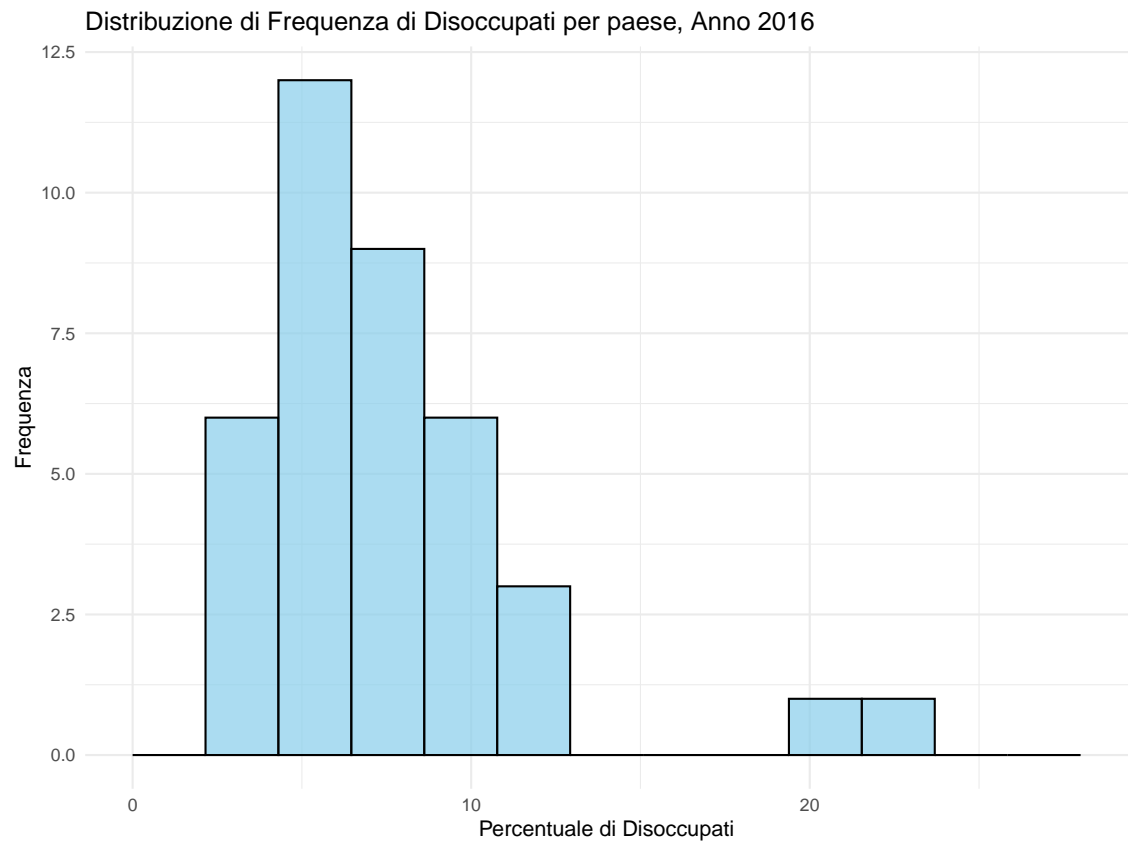


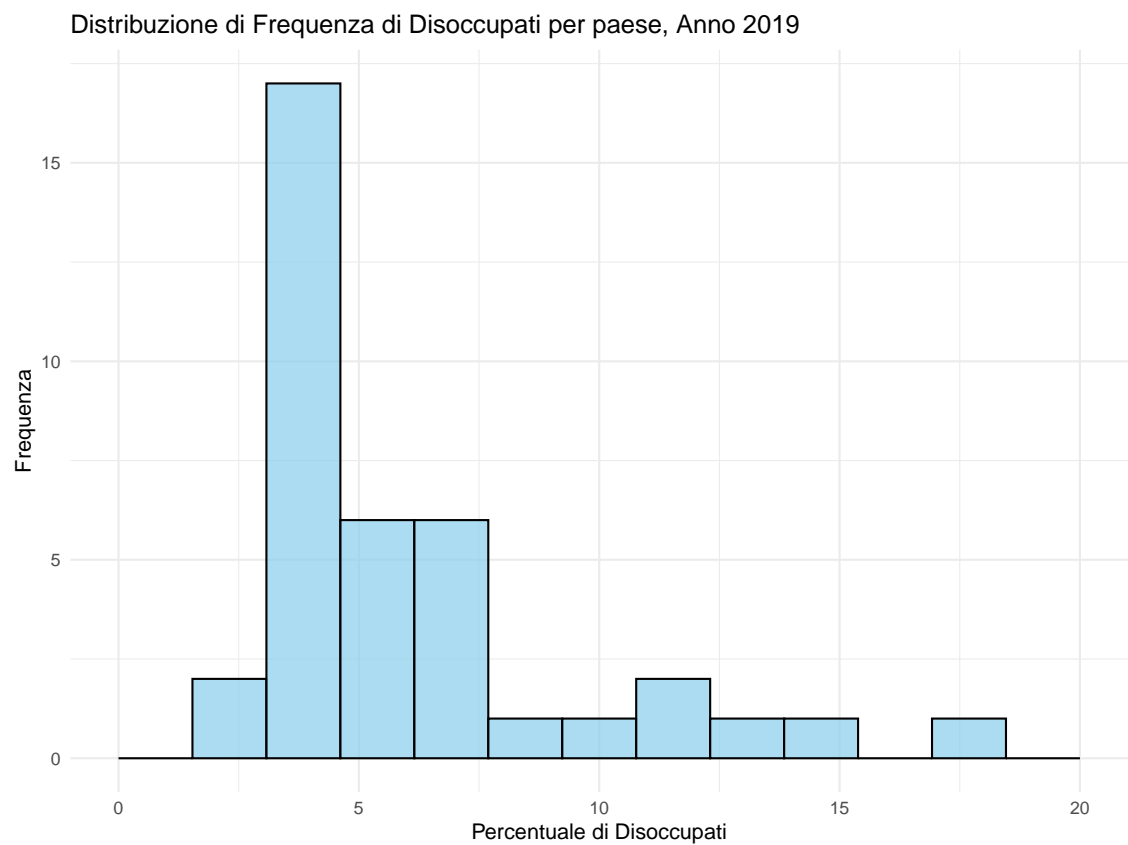
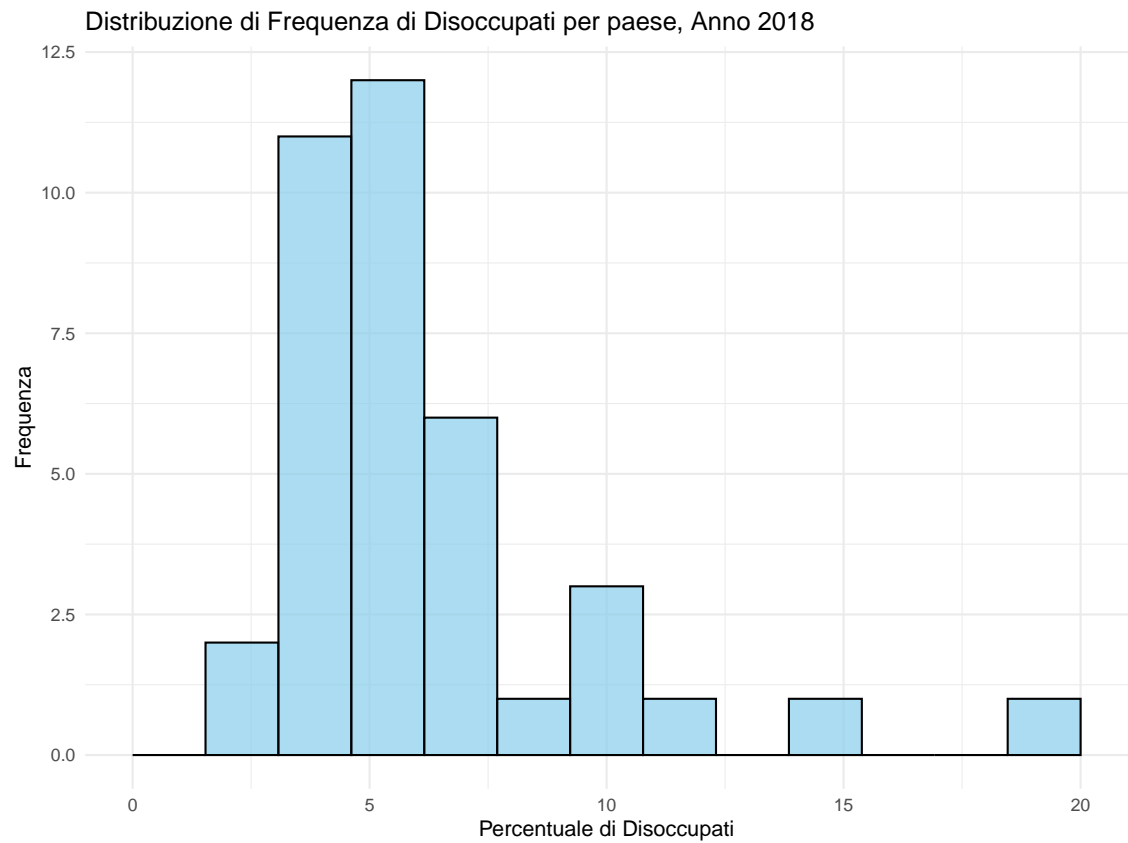
Nel 2013 sono presenti due outlier che raggiungono una percentuale di disoccupazione del 27-28%. Sono la Grecia e la Spagna, che più di tutti hanno risentito della crisi.





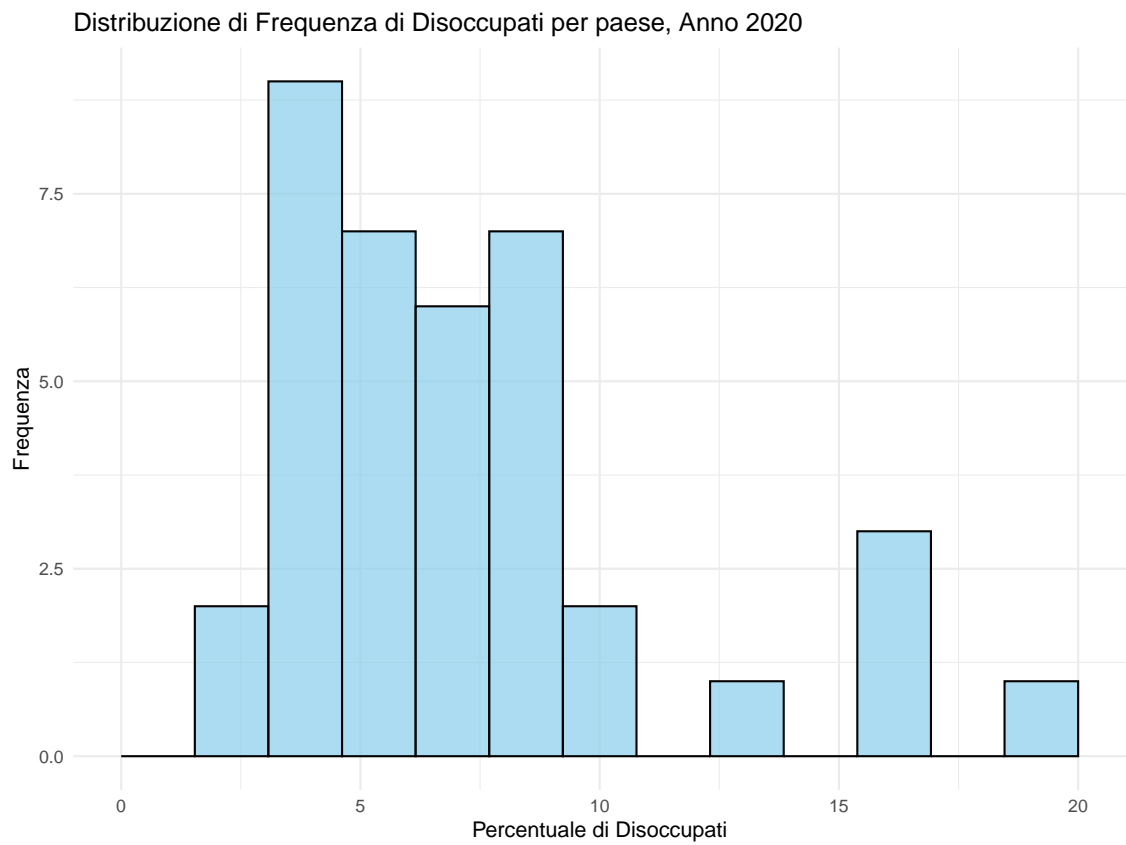
La Grecia e la Spagna, con il passare degli anni, abbassano lentamente la loro percentuale di disoccupazione.





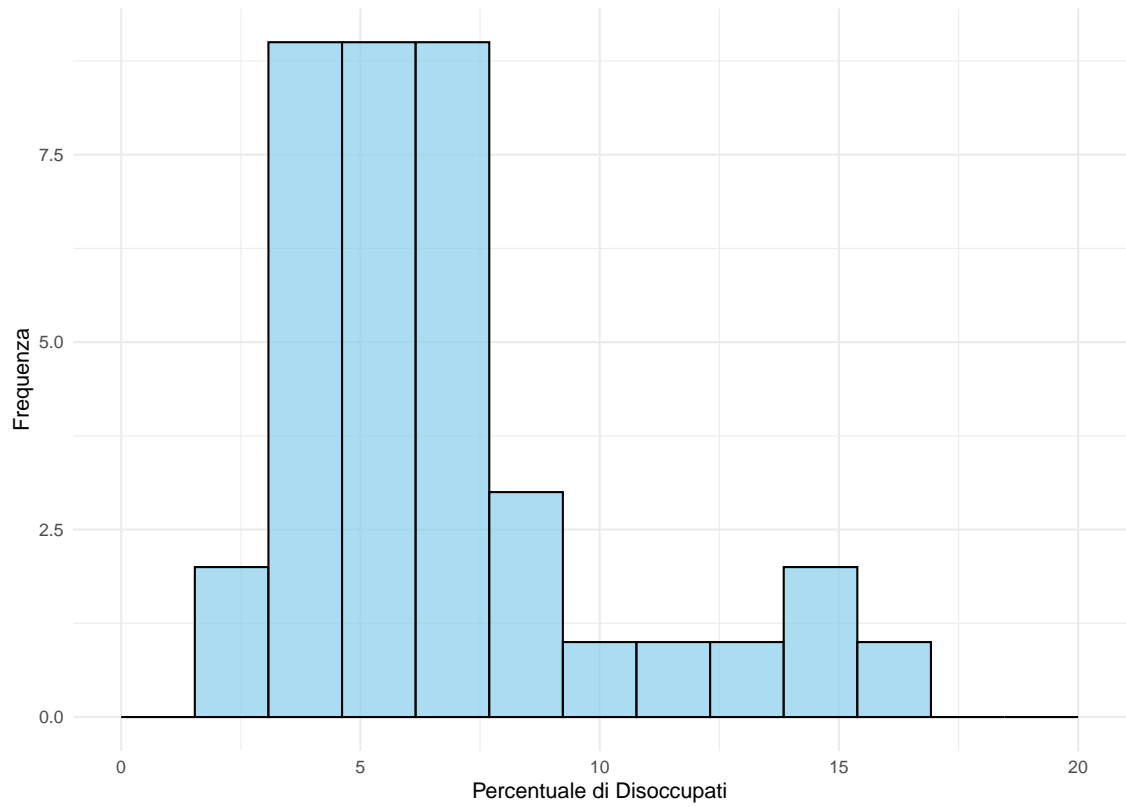


Con il passare degli anni il picco si sposta sempre più verso sinistra, diminuendo gradualmente.

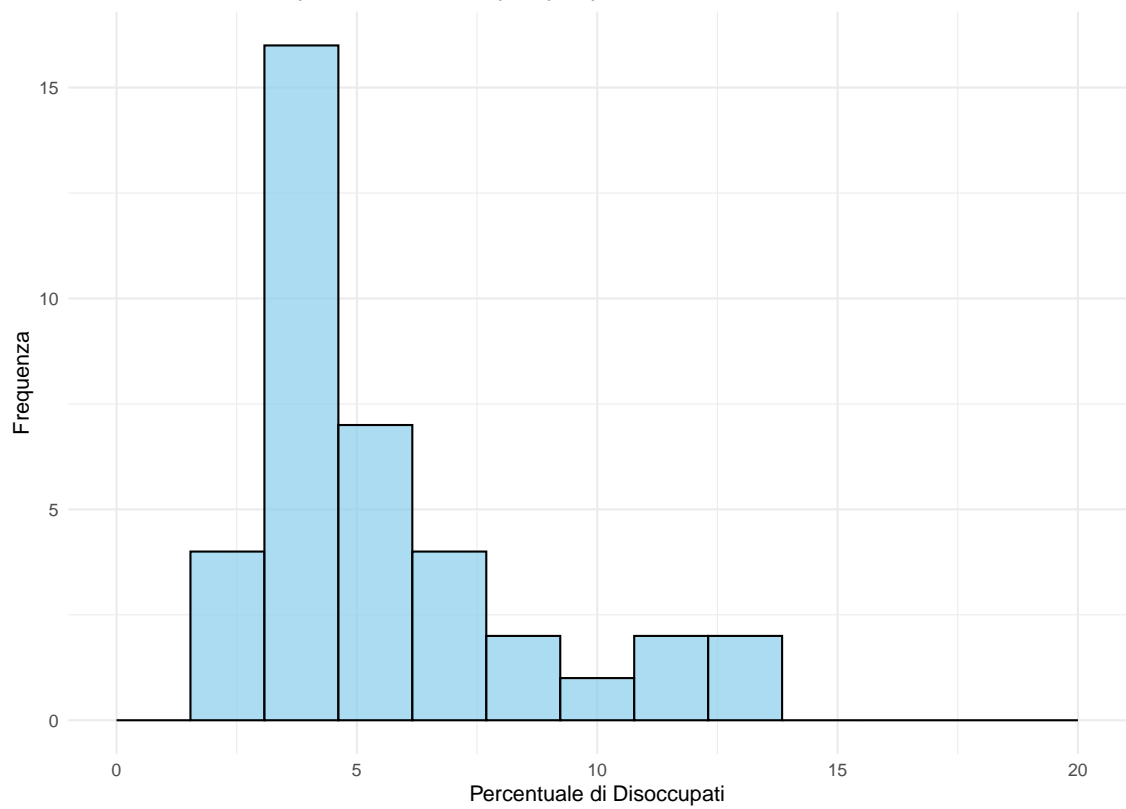


Nel 2020, a causa del COVID-19, si ha un aumento della disoccupazione. Non c'è più un picco ben definito.

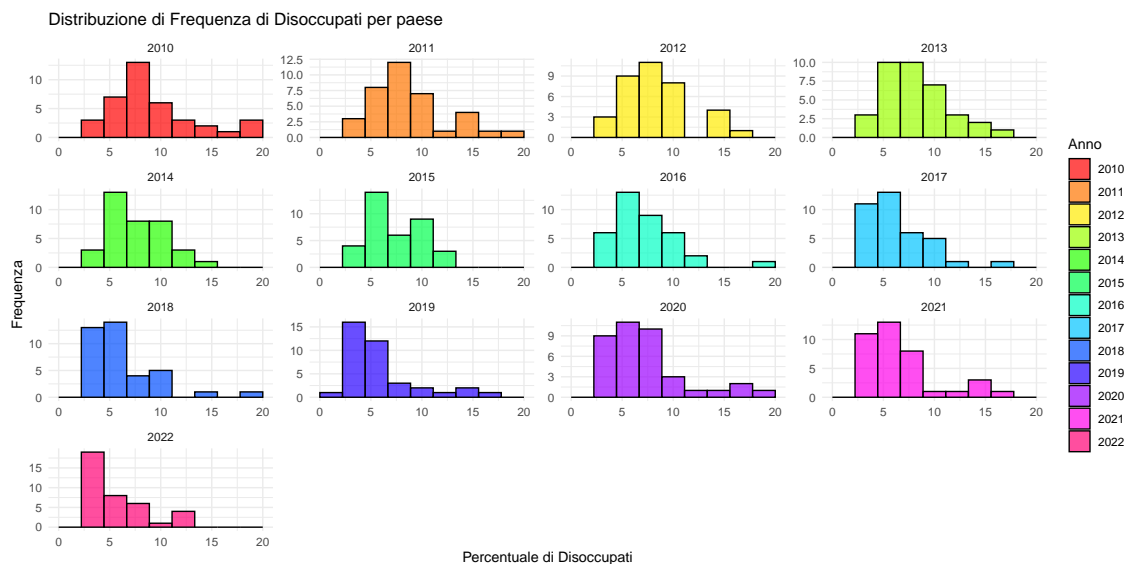
Distribuzione di Frequenza di Disoccupati per paese, Anno 2021



Distribuzione di Frequenza di Disoccupati per paese, Anno 2022



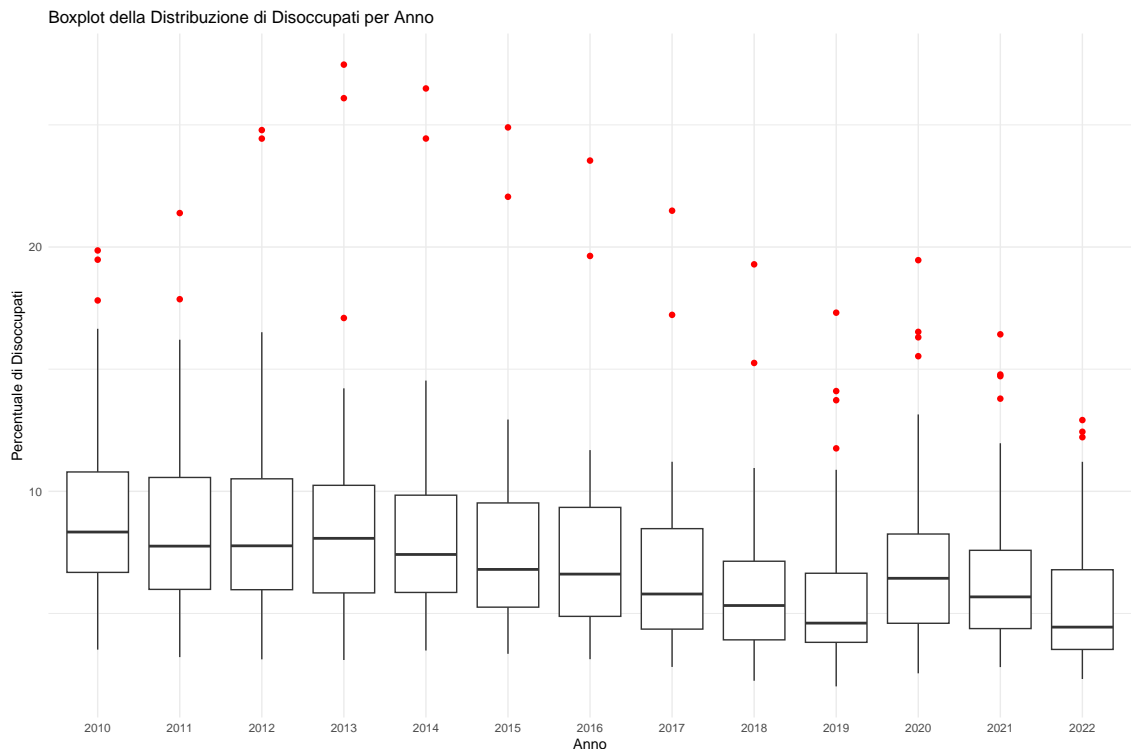
Nel 2022, come nel 2019, il picco si sposta sotto al 5%, aggirandosi attorno al 4.5%. La Grecia e la Spagna adesso hanno una percentuale dimezzata rispetto a 6 anni fa, però sono ancora le più alte in assoluto rispetto agli altri paesi, e si aggirano intorno al 12%.



Qui è possibile avere una ricapitolazione di tutti gli istogrammi nel corso degli anni.

## 4.2 Boxplot

È stato realizzato un grafico dove sull'asse delle ascisse si hanno gli anni dal 2010 al 2022, mentre sulle ordinate i valori delle percentuali di disoccupazione (rispetto alla forza lavoro). In questo modo si ha una chiara visione di come è cambiata nel tempo la mediana della percentuale di disoccupazione.



Nel corso degli anni la mediana è scesa gradualmente, tranne nel 2020 a causa della pandemia. Nel 2022 la mediana ha raggiunto il valore più basso in assoluto, il che preannuncia risultati promettenti in futuro. Nel 2013 risaltano subito all'occhio 2 outlier, che sono la Grecia e la Spagna. Questi ultimi rimangono al di fuori della soglia impostata dai baffi fino al 2022, tuttavia col passare degli anni si sono avvicinati sempre di più al baffo superiore. Non bisogna trascurare il fatto che nel 2019 gli outlier sono diventati 4 (si sono aggiunti Colombia e Costa Rica), per poi diminuire a 3 nel 2021 (la Colombia è scomparsa, ma la sua percentuale di disoccupazione ha un valore quasi pari a quello del baffo superiore).

---

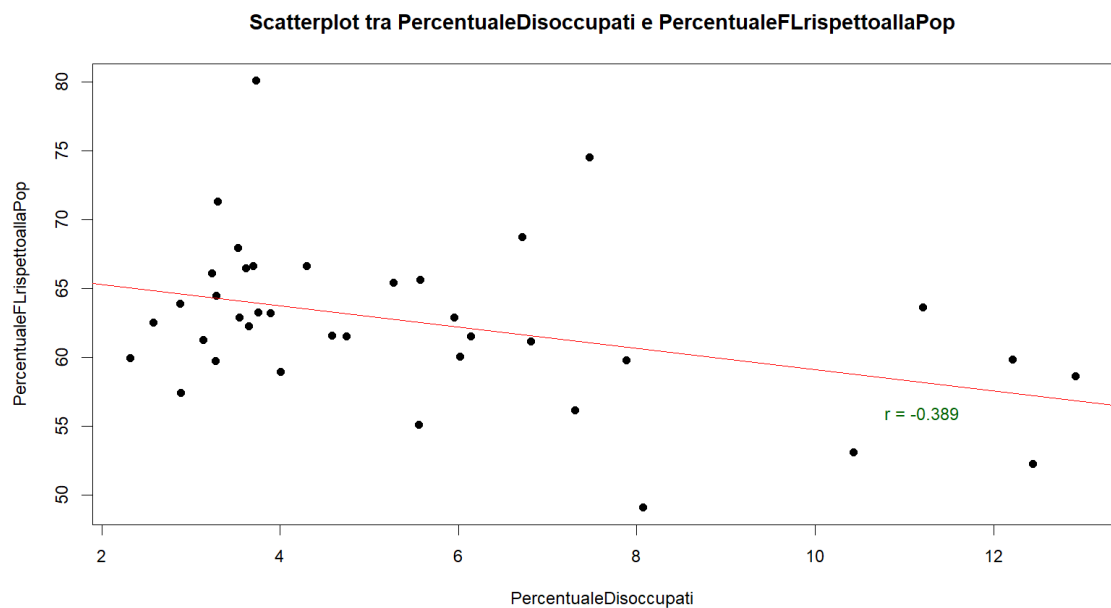
### Cluster Gerarchico

---

#### 5.1 Cluster Gerarchico

Come individui sono stati scelti i 38 paesi nel 2022, come caratteristica 1 la percentuale di disoccupati rispetto alla forza lavoro, e come caratteristica 2 la percentuale di forza lavoro rispetto alla popolazione. L'obiettivo è raggruppare i paesi in base a un'ottima organizzazione ed economia sul lavoro, o problemi simili sulla disoccupazione e sulla forza lavoro.

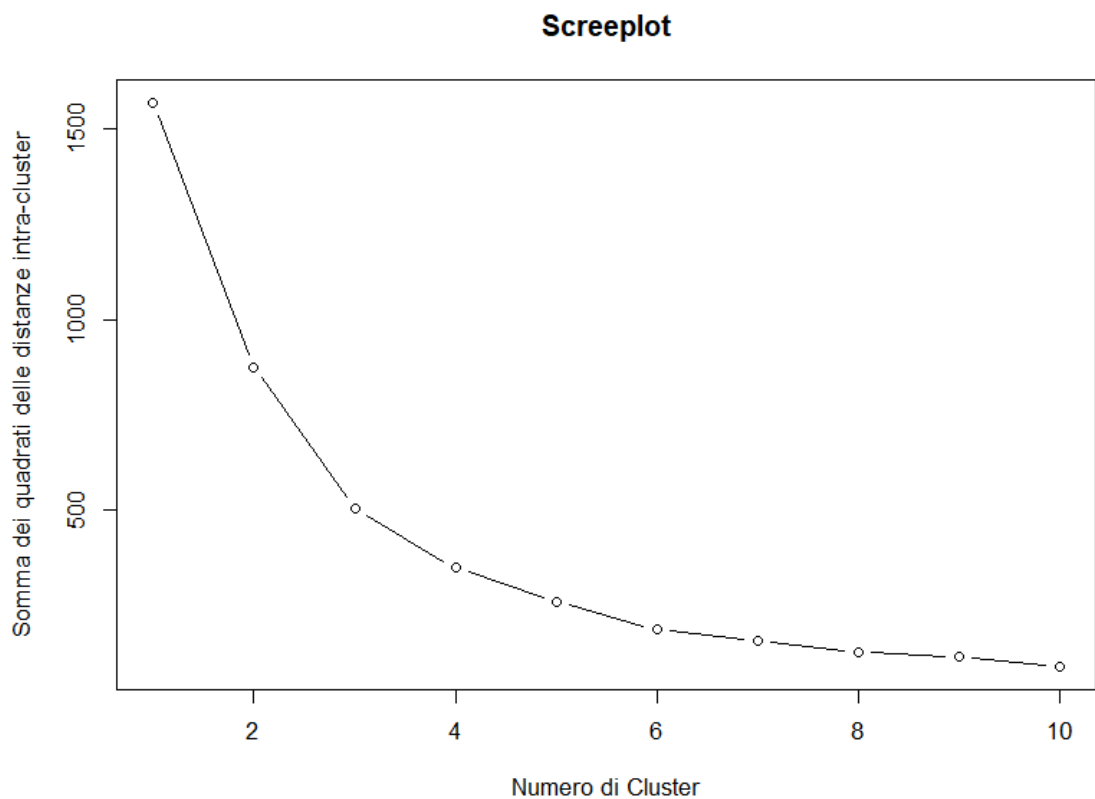
La covarianza è uguale a -6.64. Il fatto che ci sia una relazione negativa tra le due caratteristiche suggerisce che, in media, quando una delle variabili aumenta, l'altra tende a diminuire e viceversa. Di seguito è possibile visualizzare lo scatterplot delle due variabili.



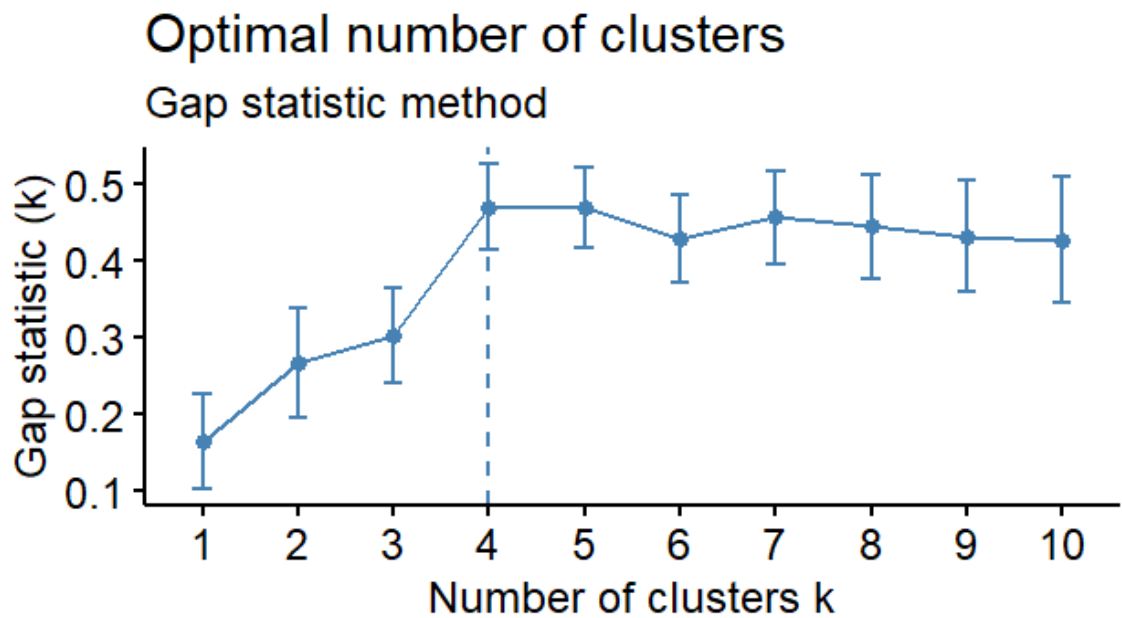
Successivamente, è stato calcolato il p-value, che è uguale a 0,0156.

Il coefficiente di correlazione negativo (-0,39) suggerisce una relazione negativa tra le variabili. In questo contesto, significa che c'è una tendenza alla diminuzione di "PercentualeFLrispettoallaPop" quando "PercentualeDisoccupati" aumenta, e viceversa. Il p-value di 0,0156 indica che il coefficiente di correlazione è statisticamente significativo a un livello di significatività del 5% (o 0,05). In altre parole, c'è evidenza statistica sufficiente per respingere l'ipotesi nulla che il coefficiente di correlazione sia uguale a zero. Un p-value inferiore a 0,05 indica che il coefficiente di correlazione è significativamente diverso da zero. Ciò significa che, anche se -0.39 non è molto vicino a -1, c'è evidenza statistica sufficiente per affermare che la relazione osservata tra le variabili "PercentualeDisoccupati" e "PercentualeFLrispettoallaPop" non è dovuta al caso.

A questo punto, attraverso l'utilizzo di metodi di valutazione dei cluster si è determinato il range di numeri per cui  $k$  è il cluster migliore.

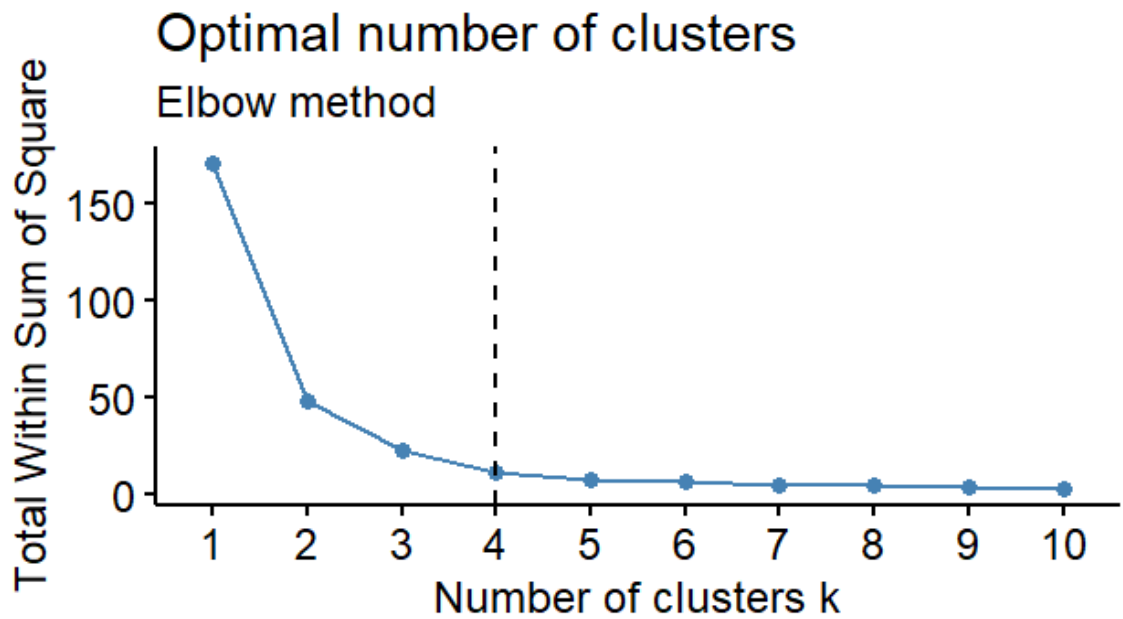


Lo Screepplot non mostra un  $k$  particolare che sia il migliore da utilizzare.

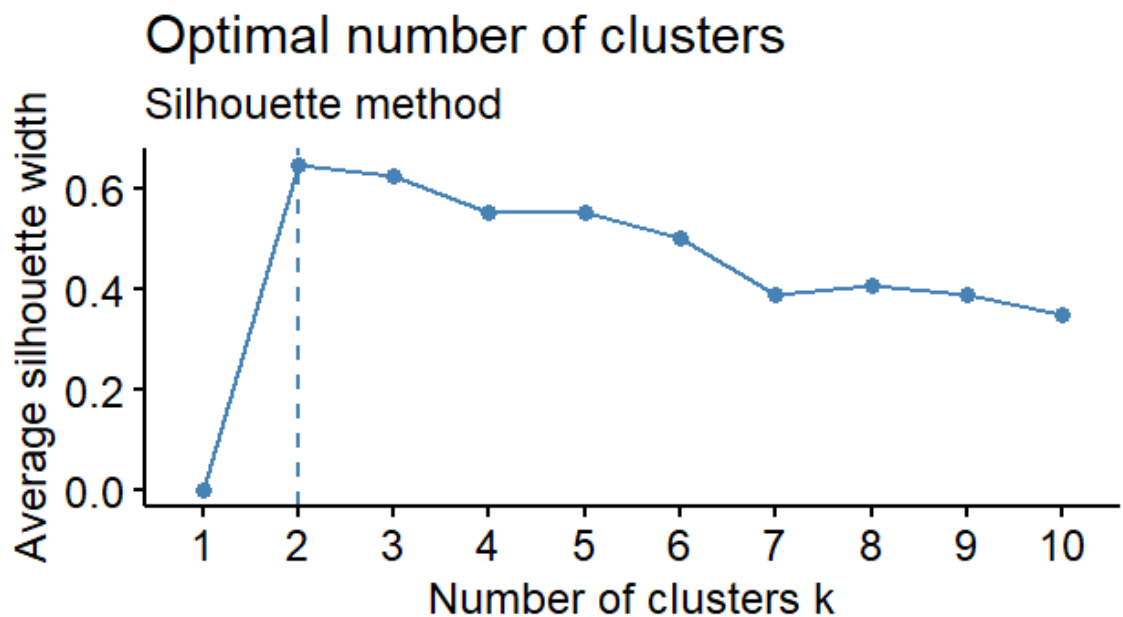


Osservando il grafico del gap sembra che il numero migliore di cluster sia pari a

4.



Osservando il grafico del gomito sembra che il numero migliore di  $k$  sia tra 2 a 4. Infatti dal numero 4 in poi si ha un plateau.

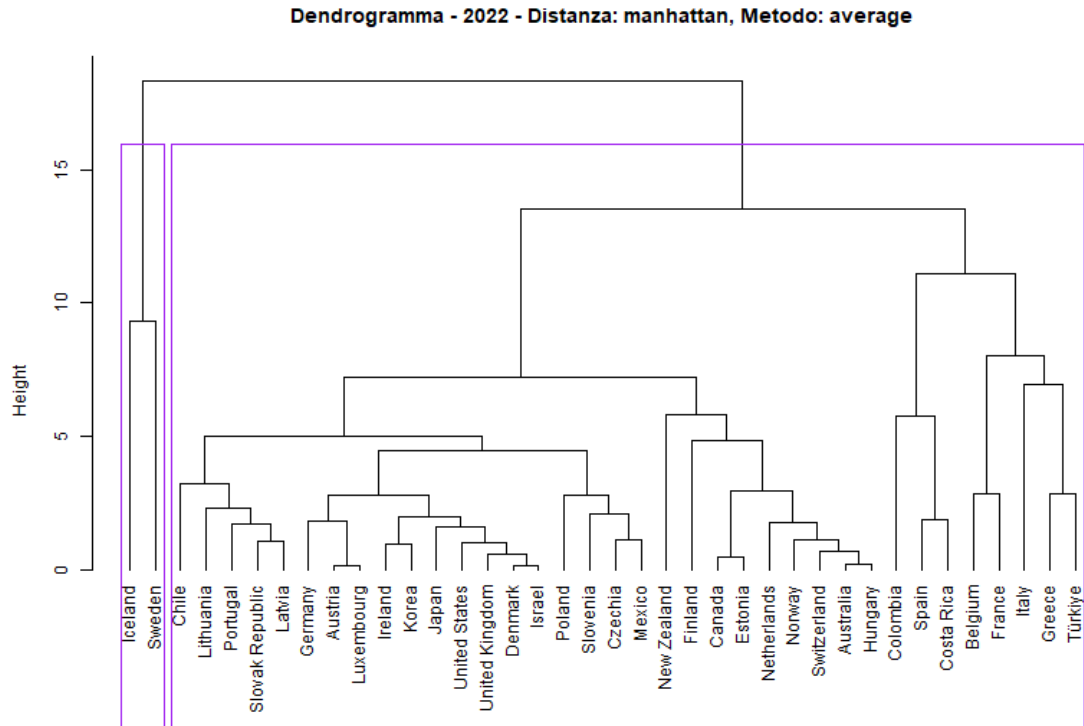


Osservando il grafico della silhouette sembra che il numero migliore di  $k$  sia tra 2 e 3.

Dopo aver considerato questi grafici, sembra che 4 sia il numero migliore di cluster, tuttavia vale la pena considerare anche  $k = 2$  e 3.



Come distanza e legame sono stati scelti Manhattan e Average, per non dare troppo peso agli outlier (più avanti saranno mostrati dendogrammi con distanze e legami diversi).



La scala delle distanze va da 0 a 18. Per  $k = 2$ , ponendo il taglio uguale a 16, si hanno 2 cluster.

Dopo vari calcoli, si ha che:

- Misure di non omogeneità statistica per il cluster 1: 56.63408
- Misure di non omogeneità statistica per il cluster 2: 2.637755
- Misura di non omogeneità interna ai cluster (within): 132.2031
- Misura di non omogeneità tra i cluster (between): 50.66151

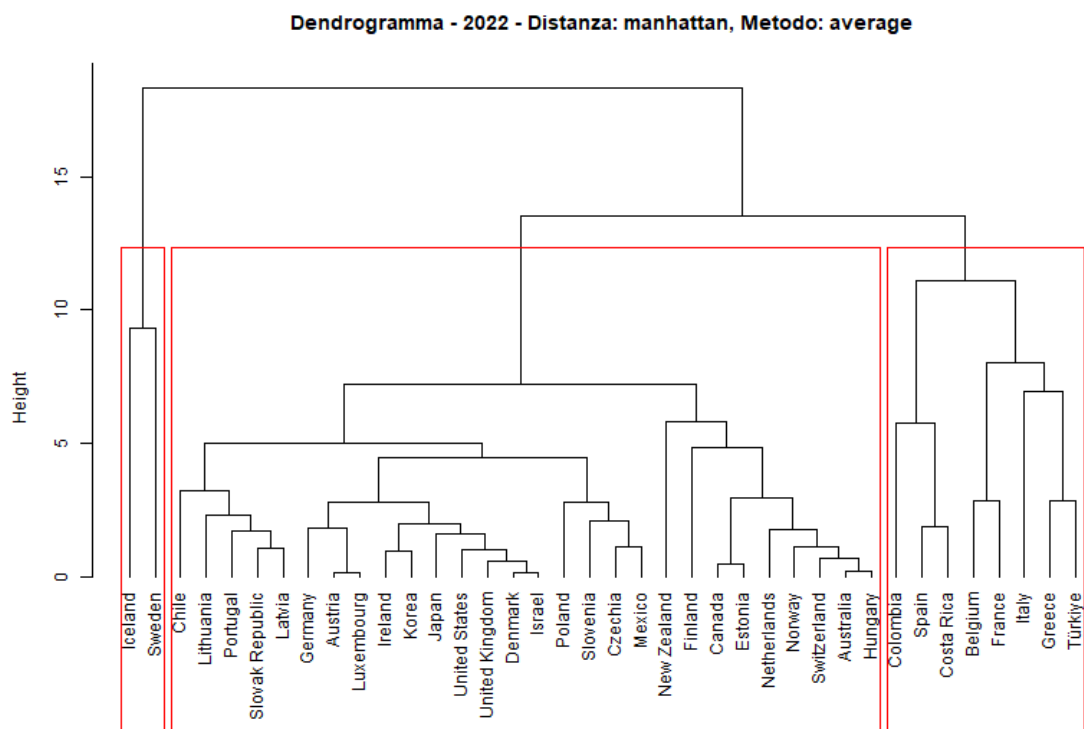
In effetti notiamo che la misura di non omogeneità statistica per il cluster a destra è abbastanza alta. Infatti, se osserviamo il dendrogramma, notiamo che il primo cluster contiene 2 paesi, mentre il secondo cluster contiene i rimanenti 36 paesi.

Facendo un'analisi sui paesi di ciascun gruppo, si ha che:

Il primo gruppo contiene Islanda e Svezia, che sono i paesi che hanno in assoluto la disoccupazione più bassa e la forza lavoro più alta. Sono quindi paesi che hanno un'ottima economia e politica sul lavoro, forza lavoro e produttività del lavoro.

Il secondo gruppo contiene sia paesi con disoccupazione media che alta, e sia paesi con forza lavoro bassa che alta.

Poiché il secondo cluster contiene paesi troppo diversi tra loro, bisogna proseguire con  $k = 3$ .



Per  $k=3$ :

- Misure di non omogeneità statistica per il cluster 1: 40.81594
- Misure di non omogeneità statistica per il cluster 2: 2.637755
- Misure di non omogeneità statistica per il cluster 3: 0.5756603
- Misura di non omogeneità interna ai cluster (within): 112.8368
- Misura di non omogeneità tra i cluster (between): 63.86406

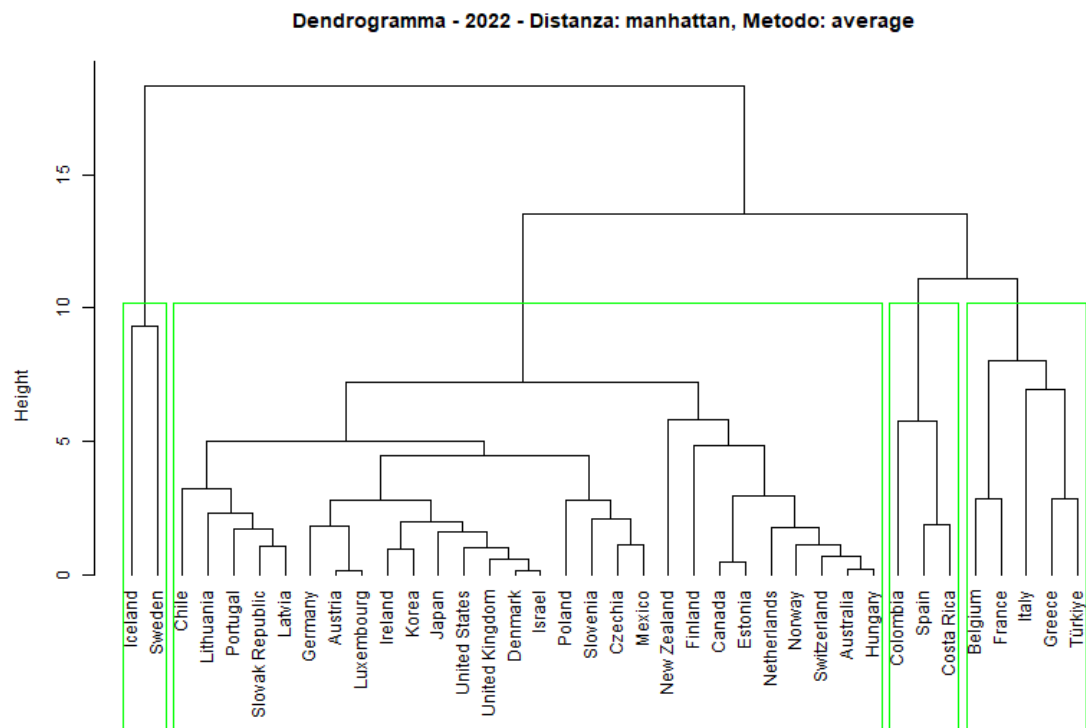
Abbiamo diminuito la distanza inter-cluster e aumentato la distanza intra-cluster.

Analizziamo i gruppi generati dal cluster. Il primo gruppo contiene, come prima, Islanda e Svezia.

Il secondo gruppo contiene tutti i paesi che hanno disoccupazione e forza lavoro nella norma. Sono quindi paesi che non si distinguono particolarmente.

Il terzo gruppo presenta i paesi con delle problematiche. Tuttavia non viene fatta una differenza in base alle problematiche.

Come vedremo nel prossimo grafico, con  $k=4$  si avrà anche una distinzione in base a questa differenza.



Per  $k=4$ ,

- Misure di non omogeneità statistica per il cluster 1: 3.574123
- Misure di non omogeneità statistica per il cluster 2: 18.15647
- Misure di non omogeneità statistica per il cluster 3: 2.637755
- Misure di non omogeneità statistica per il cluster 4: 0.5756603
- Misura di non omogeneità interna ai cluster (within): 99.63428

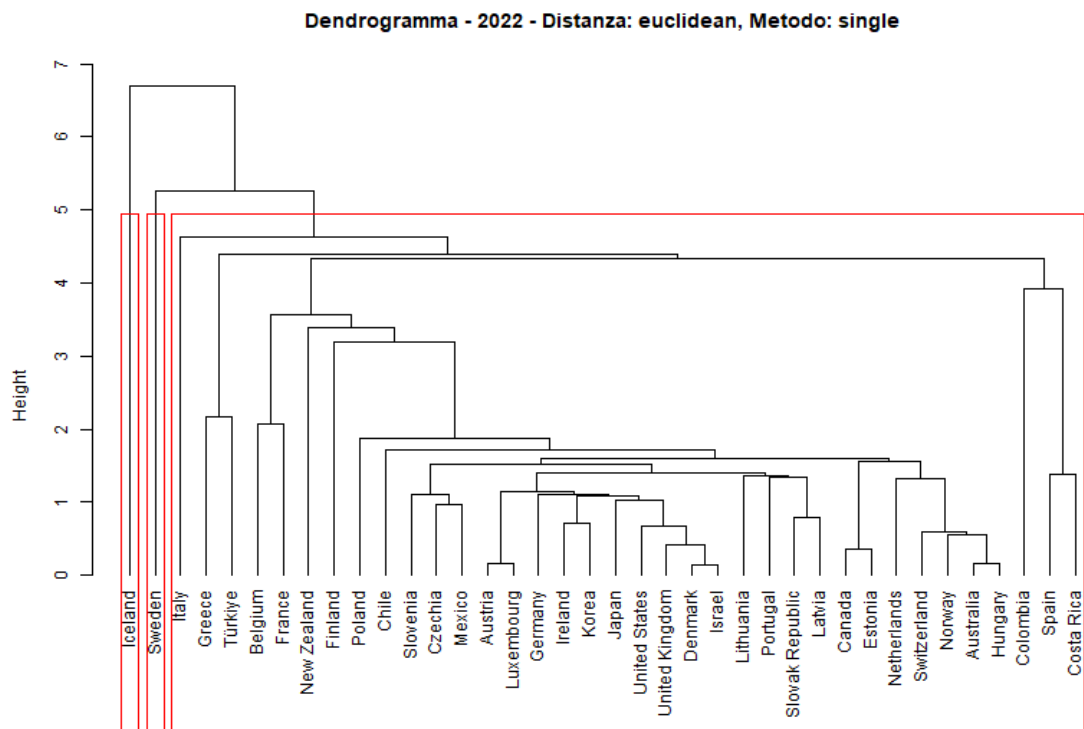
- Misura di non omogeneità tra i cluster (between): 75.63284

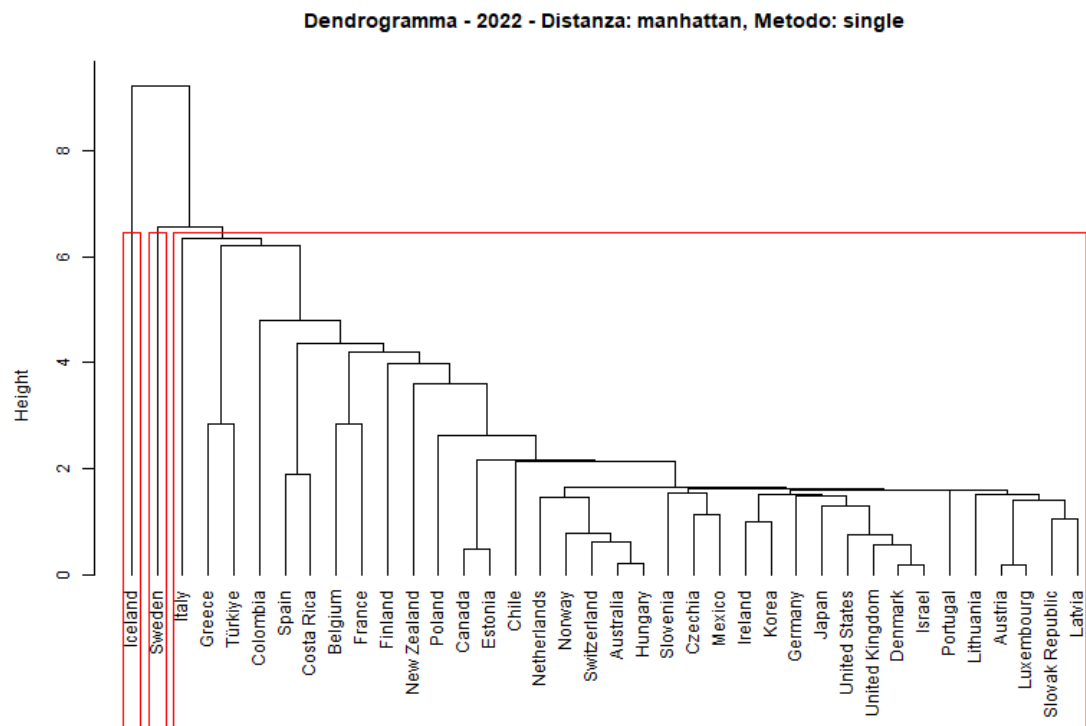
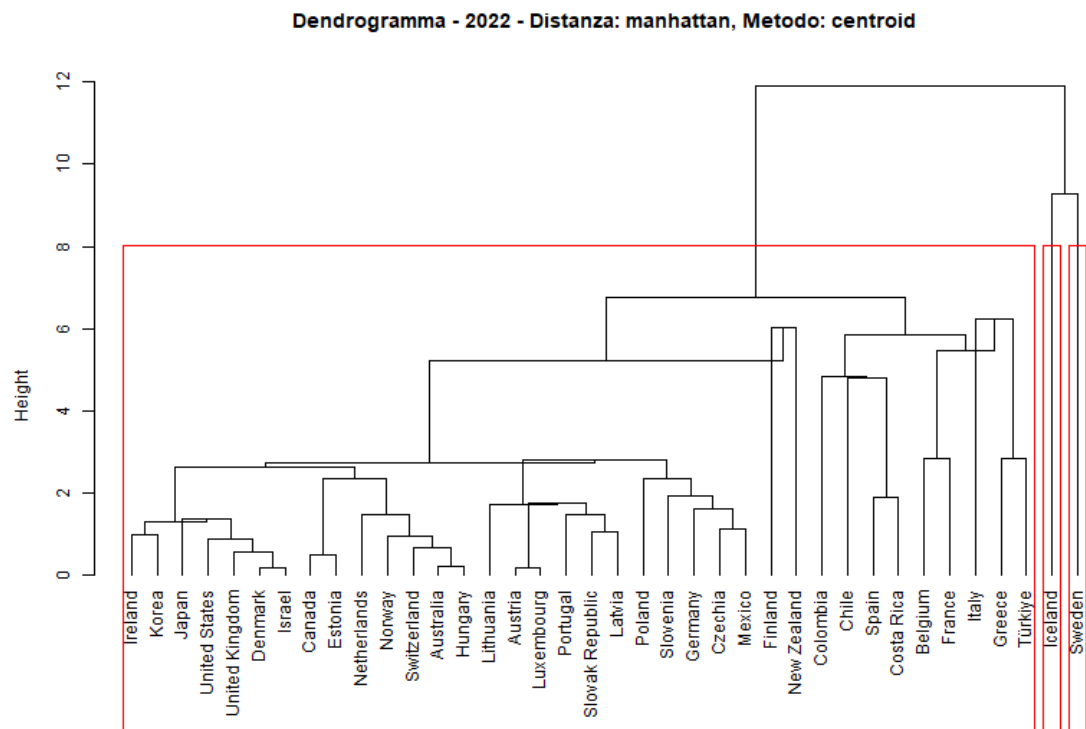
Notiamo che, rispetto a  $k=3$ , la distanza inter-cluster è diminuita e la distanza intra-cluster è aumentata. Inoltre, le misure di non omogeneità statistica interne ai singoli cluster sono diminuite.

Stavolta, il terzo gruppo contiene i paesi che hanno la disoccupazione più alta in assoluto, mentre il quarto gruppo contiene i paesi con la forza lavoro più bassa in assoluto, che sono due problematiche ben diverse.

Quindi si è scelto  $k = 4$ .

Di seguito sono riportati i cluster (anche con distanza euclidea) con legame singolo o centroide, per evidenziare quanto la presenza di outlier influenzi i cluster, causando la formazione di cluster costituiti da un singolo elemento. Infatti alla fine è stato scelto il legame medio.





## 5.2 Analisi sui cluster

Tramite il cluster gerarchico sono stati ricavati i seguenti gruppi:

1. Iceland, Sweden;
2. Chile, Lithuania, Portugal, Slovak Republic, Latvia, Germany, Austria, Luxembourg, Ireland, Korea, Japan, United States, United Kingdom, Denmark, Israel, Poland, Slovenia, Czechia, Mexico, New Zealand, Finland, Canada, Estonia, Netherlands, Norway, Switzerland, Austria, Hungary;
3. Colombia, Spain, Costa Rica;
4. Belgium, France, Italy, Turkiye.

Per quanto riguarda il primo gruppo, l'Islanda e la Svezia hanno entrambe una bassissima disoccupazione e un'altissima forza lavoro. Sono due paesi da cui si dovrebbe prendere il buon esempio. Entrambi i paesi adottano il modello svedese; Il "modello svedese" si riferisce a un sistema economico e sociale (che include diversi paesi tra cui Islanda e Svezia) basato su un mix di libero mercato e intervento statale, caratterizzato da un'elevata protezione sociale, elevati livelli di tassazione, un forte sostegno al benessere dei lavoratori e una collaborazione tra governo, sindacati e imprese per promuovere la stabilità economica e la giustizia sociale [28].

La Svezia è uno dei pochi paesi al mondo in cui è effettivamente possibile raggiungere un equilibrio tra vita lavorativa e vita privata e trascorrere del tempo di qualità con la famiglia, gli amici o dedicarsi a hobby. Una settimana lavorativa di 40 ore con orari flessibili, 480 giorni di congedo parentale retribuito per figlio e cinque settimane di ferie pagate all'anno: questi sono solo alcuni dei benefici stabiliti dalla legge svedese che consentono alle persone impiegate in Svezia di dedicare una quantità significativa di tempo ad attività al di fuori del lavoro.

Le settimane lavorative in Svezia di solito iniziano alle 8:30 o alle 9 del mattino e durano fino alle 17 dal lunedì al venerdì. Alle 17 gli uffici si svuotano e le persone tornano a casa per trascorrere del tempo con le loro famiglie. I manager non si aspettano né spingono i dipendenti a fare straordinari. Anzi, lavorare oltre l'orario previsto per giorni consecutivi ti farà avere una conversazione con un manager preoccupato per

la tua salute mentale e livelli di stress e ti aiuterà a trovare soluzioni per ridurre il tuo carico di lavoro [29].

La settimana lavorativa svedese differisce anche dalla maggior parte degli altri paesi in quanto offre orari flessibili, conosciuti in Svezia come "flextid" [30]. Flextid dà ai dipendenti il diritto di decidere come desiderano completare le loro 40 ore settimanali. Flextid può, ad esempio, significare che se si lavora oltre l'orario un giorno, si possono lavorare meno ore il giorno successivo in modo che il totale delle ore lavorate sia pari a 40 alla fine della settimana.

Il diritto di essere flessibili con gli orari di lavoro significa anche che i datori di lavoro spesso richiedono ai dipendenti di essere presenti fisicamente al lavoro solo dalle 9 alle 16, o anche dalle 10 alle 15. Manager e colleghi non reagiranno se qualcuno lascia la scrivania alle 15 per andare a prendere i figli a scuola, ad esempio. È considerato perfettamente normale. I manager comprendono le esigenze della vita familiare e si fidano dei dipendenti per completare le ore di lavoro da casa o nei giorni successivi. Tuttavia, il lavoro flessibile può comportare alcuni svantaggi per le aziende, come costi aggiuntivi per gestire programmi di lavoro flessibile e un potenziale aumento delle ore straordinarie se i dipendenti non gestiscono correttamente il proprio tempo lavorativo flessibile. Quindi è un metodo sicuramente efficiente, ma che ha ancora bisogno di alcuni miglioramenti.

La forza lavoro in Svezia e Islanda è altissima, e questo è dovuto a diversi fattori; uno dei motivi appena affrontati è il clima favorevole di lavoro. Inoltre, il sistema educativo svedese è ben sviluppato e accessibile, il che consente a un'ampia parte della popolazione di acquisire le competenze necessarie per partecipare attivamente al mercato del lavoro. Infine, la Svezia ha adottato politiche attive e inclusive del mercato del lavoro, come politiche di riqualificazione e attivazione, che aiutano a ridurre la disoccupazione di lungo termine e ad aumentare il coinvolgimento della forza lavoro.

Si potrebbe fare un passo avanti e ipotizzare che, sulla base dei dati osservati sui paesi che adottano la politica del modello svedese, non è necessariamente vero che le nuove generazioni non vogliano lavorare [31]. Infatti, rendendo l'ambiente di lavoro più flessibile e tutelato, si può osservare come la percentuale di disoccupati scenda drasticamente.

Per quanto riguarda il secondo gruppo, esso è formato da paesi che sono nella media e non presentano caratteristiche particolari.

Il terzo gruppo presenta paesi con la percentuale di disoccupazione più alta in assoluto. Sono 3 paesi abbastanza diversi tra loro. Mentre la Costa Rica ha solo 5 milioni di abitanti, la Spagna e la Colombia hanno intorno ai 50 milioni di abitanti. Solo la Spagna fa parte dell'Europa, mentre le altre due dell'America Latina. Questi tre paesi non si sono ancora ripresi completamente dalla crisi del 2008, inoltre la Colombia e la Costa Rica hanno un sistema educativo inefficiente o poco adattato alle esigenze del mercato del lavoro, e ciò può creare una discrepanza tra le competenze richieste dalle imprese e quelle possedute dalla forza lavoro. In Costa Rica la situazione della disoccupazione sembra essere correlata a una discrepanza tra domanda e offerta di lavoro dovuta ai cambiamenti strutturali nell'economia, sia in termini di attività economica che di settore istituzionale [32]. Inoltre, avendo osservato in precedenza le serie temporali di Colombia e Costa Rica, è possibile notare come tali paesi abbiano avuto una crescita della forza lavoro lenta, incostante e con diverse ricadute. Un dato promettente è che tutti e tre i paesi hanno avuto un calo della disoccupazione sia nel 2021 che nel 2022, tuttavia sarebbe necessario un calo più ripido per poter ritornare a una percentuale di disoccupazione nella media.



Il quarto gruppo è formato dai paesi con la forza lavoro più bassa in assoluto. E al primo posto, con una percentuale di forza lavoro del 49% (meno di metà della popolazione) c'è proprio l'Italia. Ciò significa che c'è una parte consistente della popolazione che non cerca lavoro e rimane volontariamente disoccupata. Bisogna dire che questo valore è influenzato anche dalla presenza di molte persone anziane in Italia, come è stato possibile vedere precedentemente dai barplot dell'età. Inoltre, anche la "fuga di cervelli" influisce sulla diminuzione della forza lavoro [33]. Le persone potrebbero essere disposte a lavorare solo se il salario offerto soddisfa le loro aspettative. Se le offerte di lavoro disponibili non soddisfano tali aspettative, le persone potrebbero ritirarsi dal mercato del lavoro fino a quando non trovano un'opportunità che ritengono adeguata. Anche lo sfruttamento e la mancanza di sicurezza sul lavoro sono aspetti importanti da considerare. [34]

---

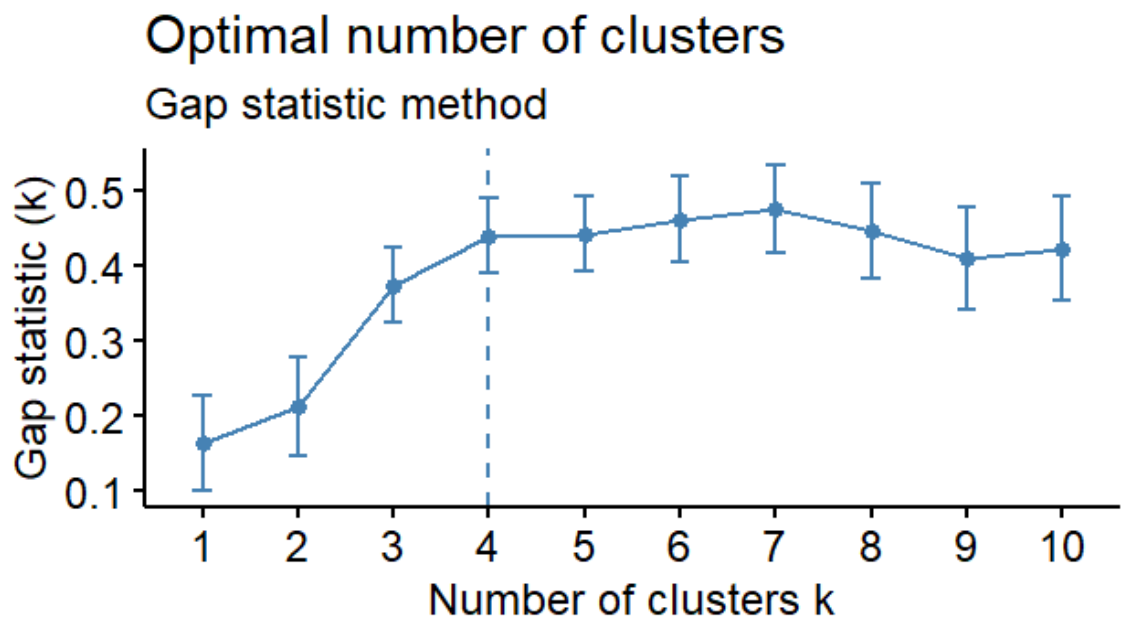
### Cluster K-MEANS

---

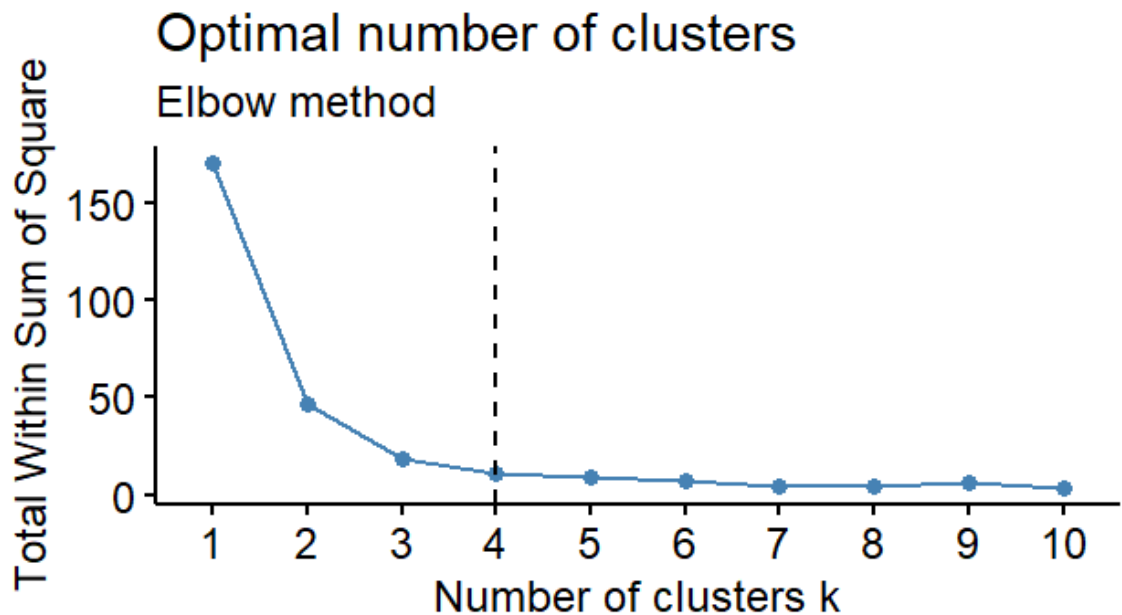
#### 6.1 Cluster K-MEANS

Come individui sono stati scelti i 38 paesi nle 2022, come caratteristica 1 la percentuale di disoccupati maschile rispetto alla forza lavoro, e come caratteristica 2 la percentuale di disoccupati femminile rispetto alla forza lavoro. L'obiettivo è raggruppare i paesi in base al tipo di problematiche che hanno sulla disparità di disoccupazione tra i due sessi.

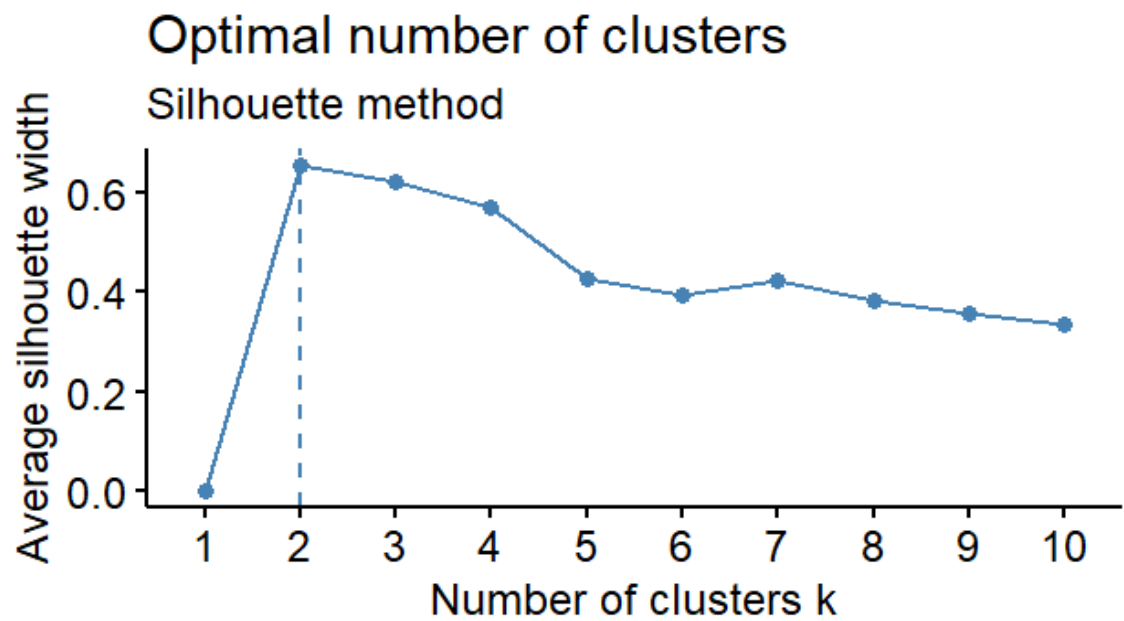
Come nel capitolo precedente, utilizziamo i metodi di valutazione dei cluster per determinare il range di numeri per cui  $k$  è il cluster migliore.



Osservando il grafico del gap sembra che il numero migliore di cluster sia attorno al 4.

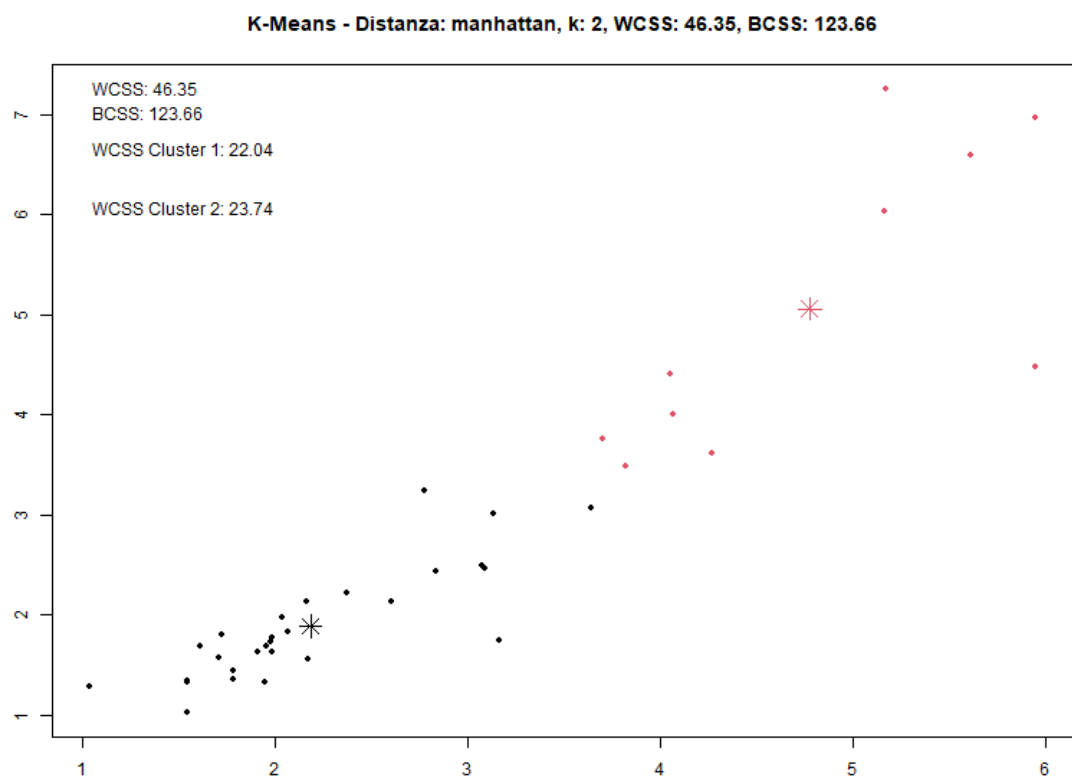


Osservando il grafico del gomito sembra che il numero migliore di  $k$  sia tra 2 a 4.



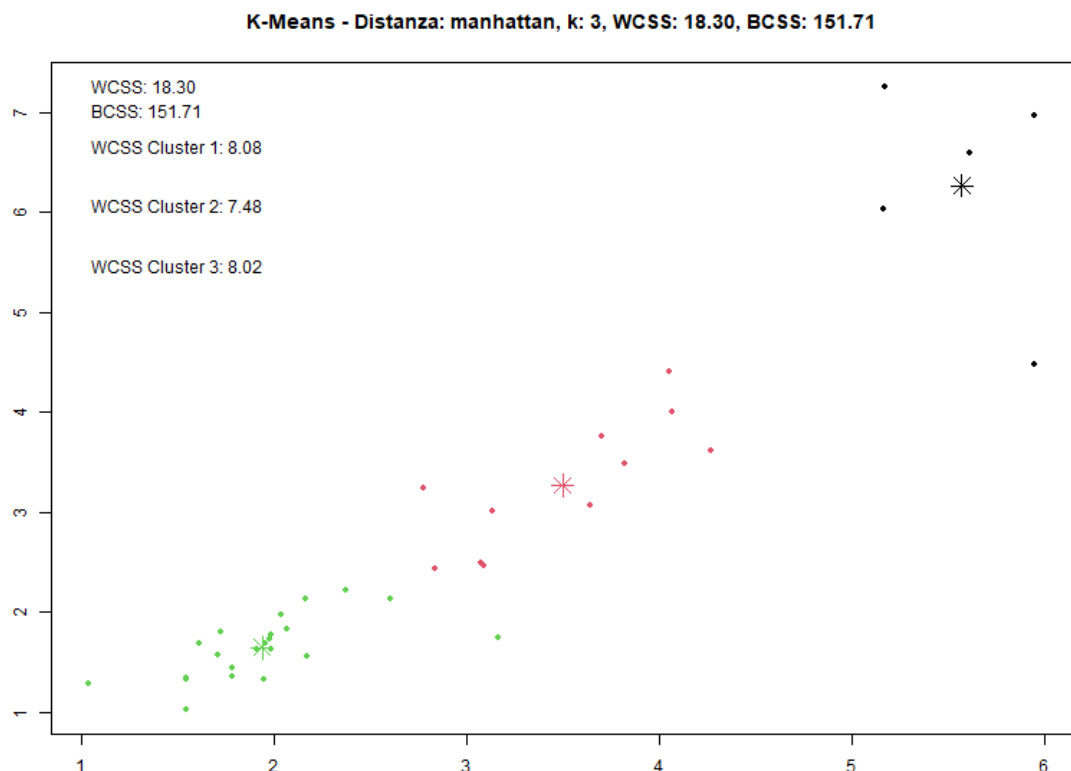
Osservando il grafico della silhouette sembra che il numero migliore di k sia tra 2 e 3.

Dopo aver valutato questi metodi, consideriamo  $k = 2, 3, 4$ .



Per ogni grafico sono riportate in alto le misure di non omogeneità intra-cluster, inter-cluster e tra i singoli cluster. Nel caso in cui  $k = 2$ , la misura di non omogeneità inter-cluster è pari a 46.35, che è più del doppio della stessa misura per  $k = 3$ . Bisogna cercare di minimizzare tale misura.

Facendo un'analisi sui due gruppi, si ha che il gruppo avente il centroide in basso a sinistra nel grafico contiene paesi con una percentuale di disoccupazione maschile/femminile medio-bassa, mentre l'altro gruppo contiene paesi con una percentuale di disoccupazione maschile/femminile medio-alta. Ciò non risulta molto utile ai fini della nostra analisi, in quanto l'obiettivo è evidenziare anche i paesi con uno squilibrio tra le percentuali di disoccupazione dei due sessi, e quindi i due cluster non sono interessanti per la nostra ricerca. Di conseguenza si è scelto di procedere con 3 cluster.

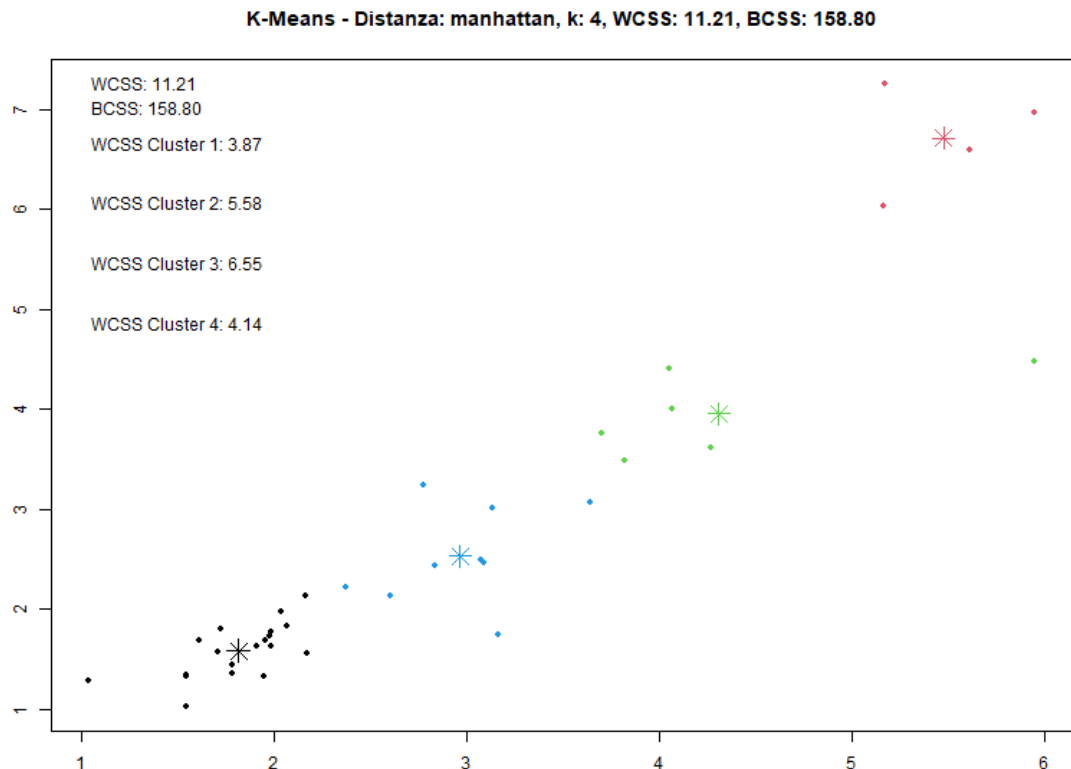


Per  $k = 3$ , le misure di non omogeneità sono migliori rispetto al caso in cui  $k = 2$ . Analizzando i 3 cluster, partendo dal basso a sinistra:

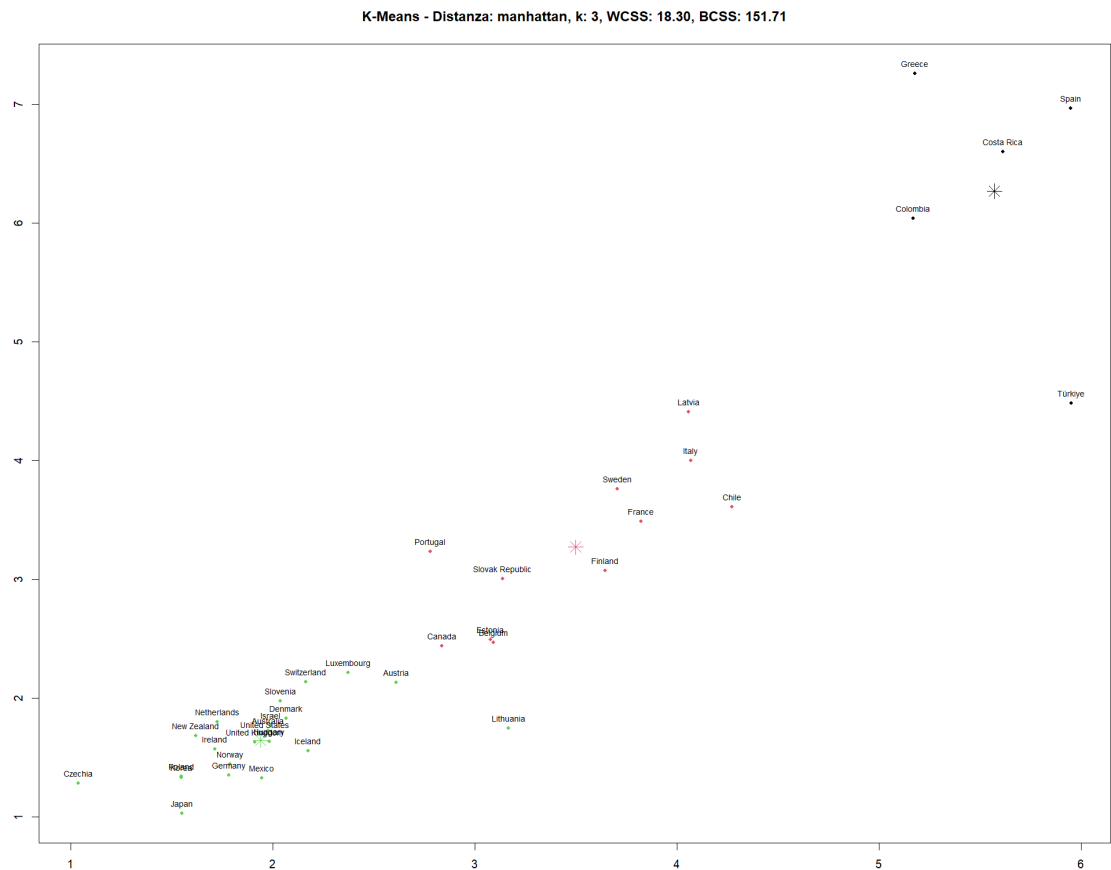
Il primo gruppo contiene i paesi con una percentuale di disoccupazione maschile/femminile bassa.

Il secondo gruppo contiene i paesi con una percentuale di disoccupazione maschile/femminile media.

Il terzo gruppo contiene paesi con una forte differenza tra percentuale di disoccupati maschile e femminile, e in cui almeno una delle due percentuali ha un valore alto. Questo perché per esserci una grande differenza, è necessario che almeno una delle due percentuali abbia un valore medio-alto.



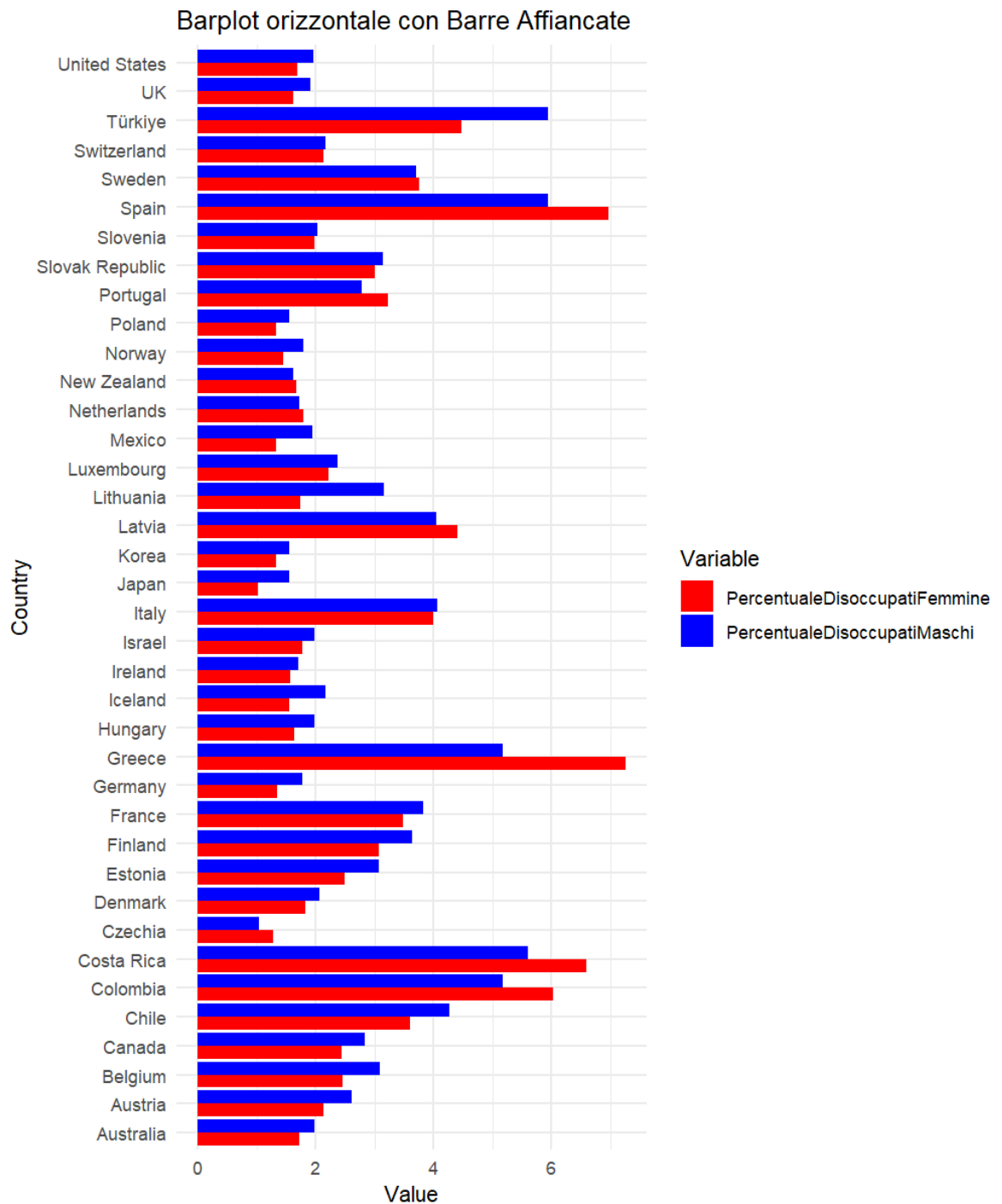
Per  $k = 4$ , si ha una ulteriore suddivisione degli elementi centrali del grafico. Dopo una analisi sulla suddivisione, i nuovi due gruppi presentano in realtà una differenza lieve, poiché il primo contiene una percentuale media di disoccupazione maschile/femminile, mentre il secondo gruppo contiene una percentuale leggermente sopra la media di disoccupazione maschile/femminile. Poiché questa non è una suddivisione rilevante per la nostra ricerca, si è scelto di utilizzare  $k = 3$ .



Qui è possibile vedere le etichette di ogni paese, che sono state utilizzate per fare analisi sui cluster.

## 6.2 Analisi sui cluster

Di seguito è riportato un barplot delle percentuali di disoccupazione maschile e femminile. I paesi appaiono in ordine alfabetico, nel caso in cui si volesse consultare le percentuali.



I 3 cluster sono i seguenti:

1. Czechia, Japan, Korea, Poland, New Zealand, Netherlands, Ireland, Norway, Germany, Mexico, United Kingdom, United States, Iceland, Australia, Iceland, Israel, Denmark, Slovenia, Switzerland, Luxembourg, Austria, Lithuania;
2. Portugal, Canada, Estonia, Belgium, Slovak Replubic, Finland, France, Sweden, Italy, Chile, Latvia;



### 3. Colombia, Greece, Turkiye, Costa Rica, Spain.

Il primo gruppo contiene i paesi che in generale hanno entrambi le percentuali di disoccupazione basse. Unica eccezione per la Lituania, che nonostante abbia percentuali nella media, ha molti meno disoccupati femminili che maschili. Tale problematica della Lituania è stata già affrontata in precedenza.

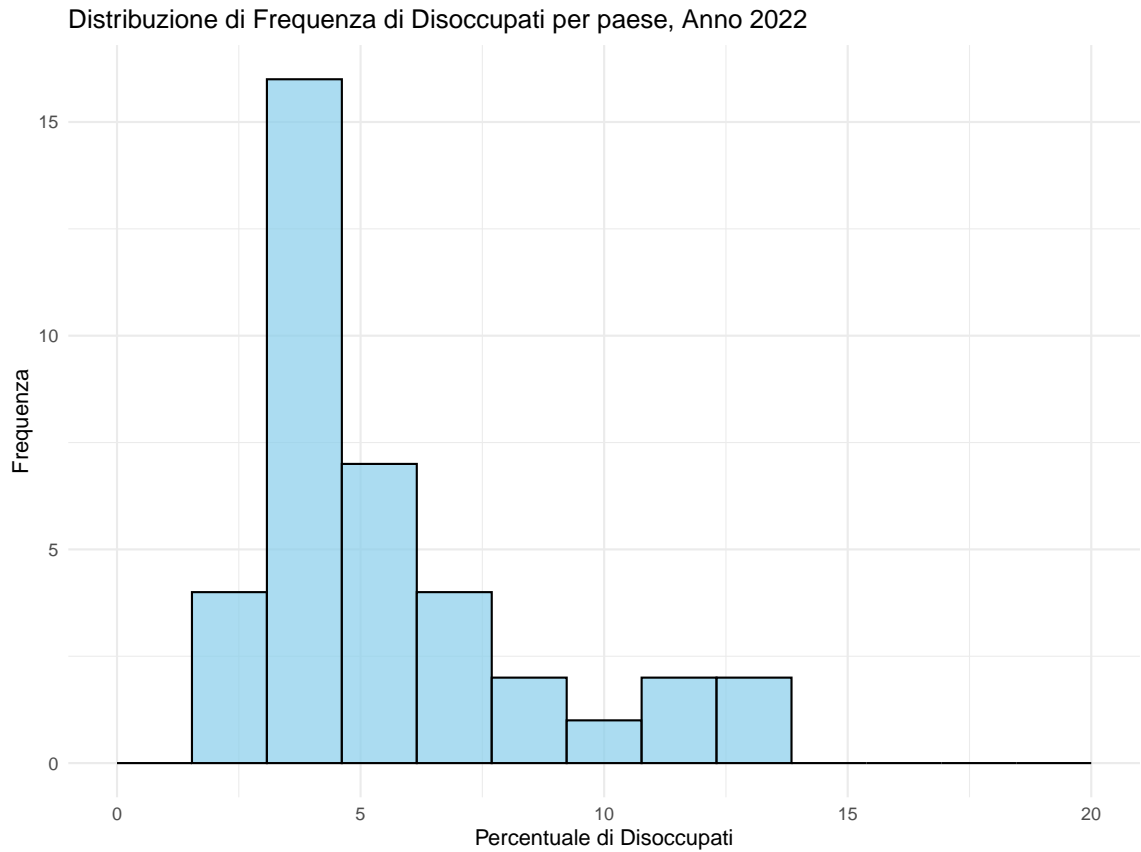
Il secondo gruppo contiene i paesi che hanno entrambe le percentuali di disoccupazione nella media o un po' sopra la media. Tuttavia per ogni paese le due percentuali sono bilanciate.

Il terzo gruppo presenta paesi che hanno una percentuale di disoccupati femminile molto più alta rispetto a quella maschile. L'unica anomalia è la Turchia. È interessante notare come nell'80% dei paesi del terzo cluster la percentuale di disoccupati femminile sia molto più alta di quella maschile. È altrettanto interessante notare come, in un gruppo di "eccezioni", sia presente a sua volta una "eccezione" come la Turchia, che ha molti più disoccupati maschi. È l'unico paese su 38 a presentare questa caratteristica.

### 7.1 Scelta della funzione di distribuzione

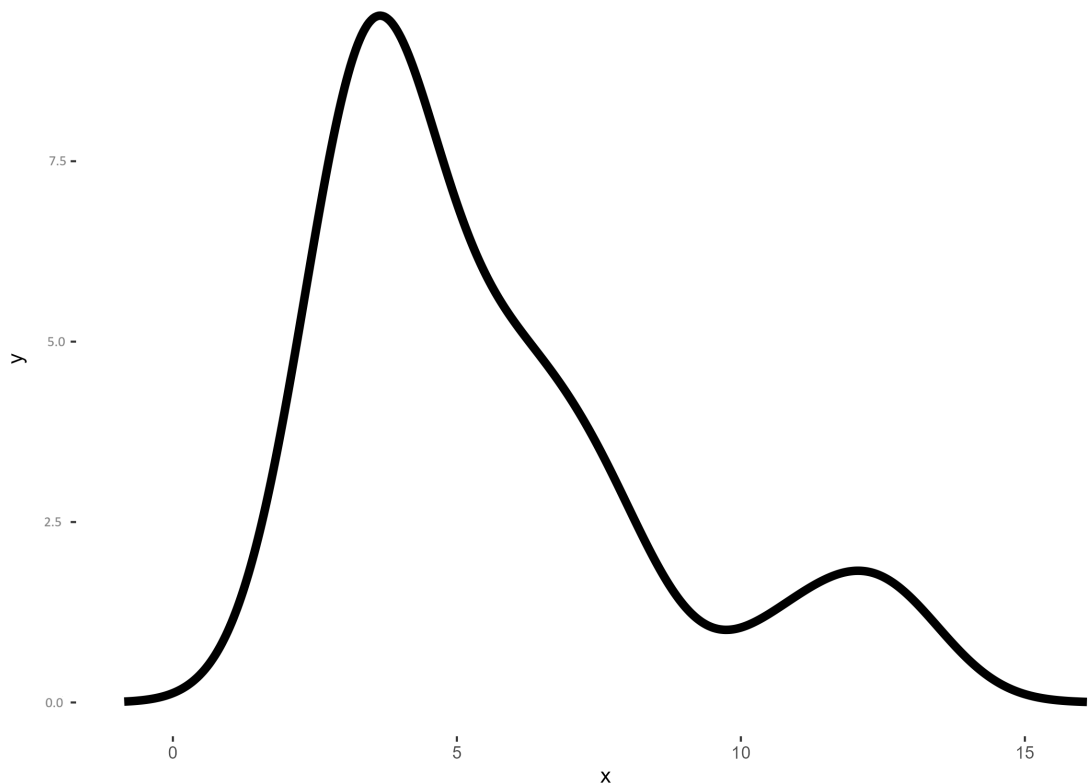
Una variabile aleatoria  $X$  è una funzione definita sullo spazio campionario  $\Omega$  che associa ad ogni evento  $E$  dell'insieme  $\Omega$  un unico numero reale. Una variabile aleatoria può essere classificata come:

- Discreta: può assumere un insieme discreto (finito o numerabile) di numeri reali.
- Continua: può assumere tutti i valori compresi in un intervallo reale.



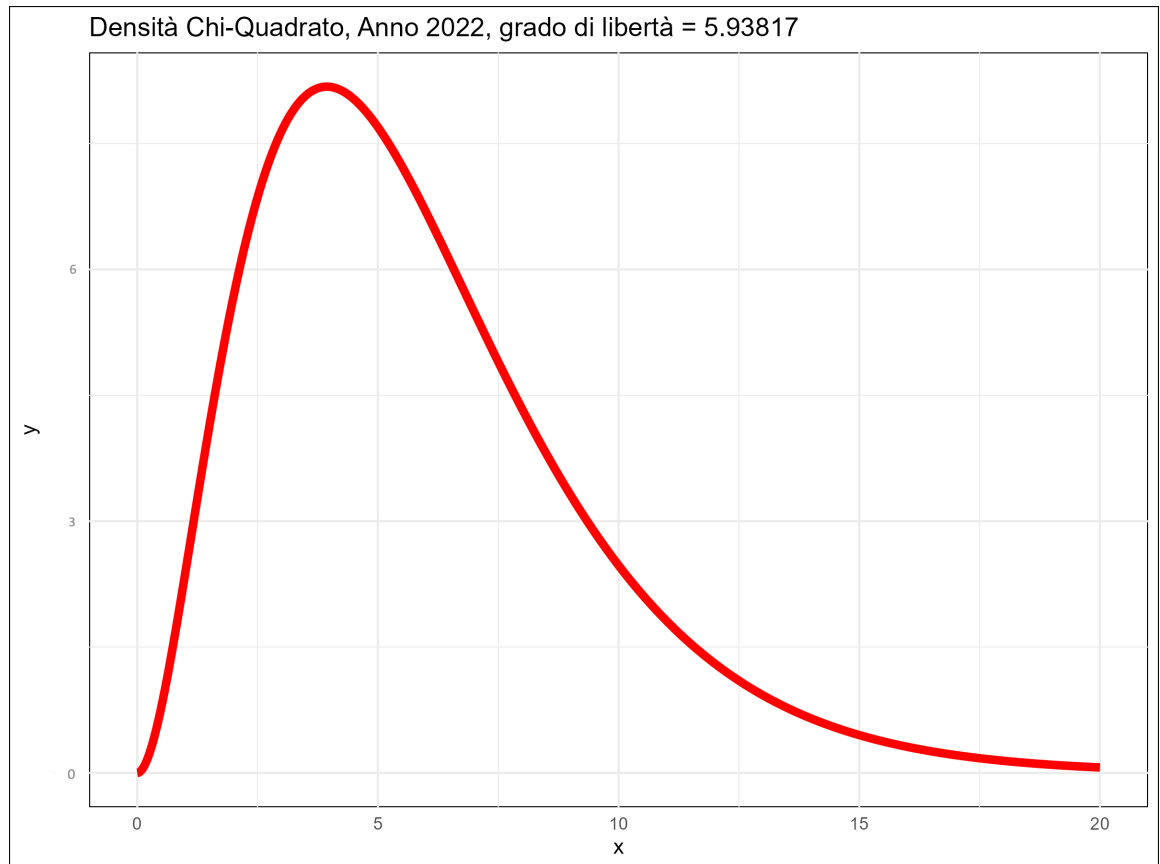
Il seguente istogramma, come già visto nel capitolo degli istogrammi, contiene la distribuzione delle frequenze delle percentuali di disoccupati (rispetto alla forza lavoro) nel 2022, per 38 paesi. Si tratta quindi di variabili aleatorie continue, poiché la percentuale può essere un numero qualsiasi da 0 a 100 (Esempio: 3.28323474). Di seguito, possiamo vedere la funzione di densità di probabilità corrispondente.

Gaussian Kernel - Frequenza Percentuale Disoccupati 2022

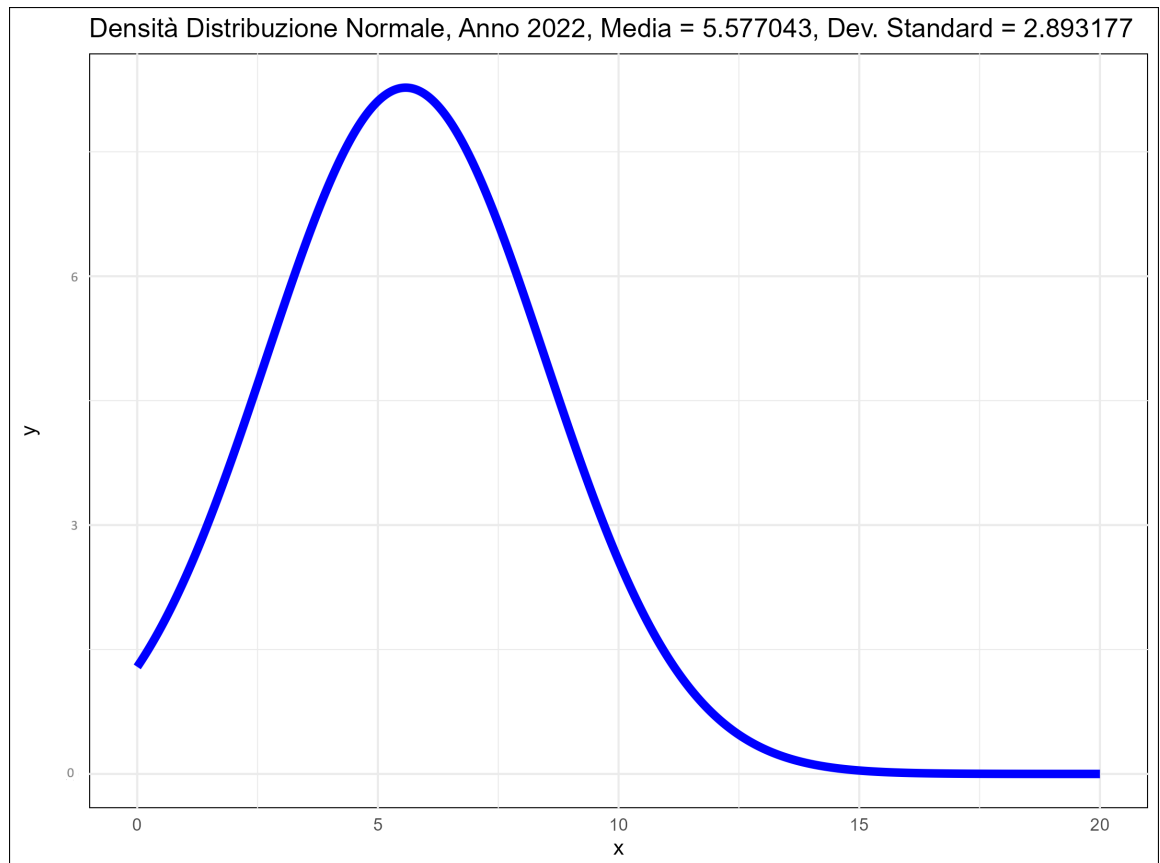


Quale funzione di distribuzione si adatta meglio al nostro caso? La distribuzione uniforme è scartata a prescindere poiché non è vero che ogni punto dell'intervallo  $[0, 100]$  ha la stessa probabilità di essere scelto. Non può essere una distribuzione esponenziale perché la frequenza delle percentuali di disoccupati non cresce esponenzialmente, e possiamo notare un picco intorno al valore 4.5. Allontanandosi da questo valore, la funzione di densità decresce. A questo punto, rimane da scegliere tra la distribuzione Chi Quadro, la distribuzione Normale e la distribuzione Student. Osservando la funzione di densità, notiamo che la curva è asimmetrica. Questa caratteristica appartiene alla distribuzione Chi Quadro, ma non alle distribuzioni Normale e Student. Come funzione di distribuzione è stata quindi scelta la distribuzione Chi Quadro. Nelle prossime sezioni saranno date ulteriori giustificazioni di questa scelta.

## 7.2 Stima puntuale e Intervalli di Confidenza



Prendiamo in considerazione la distribuzione Chi Quadrato. Utilizzando la stima puntuale, si è ricavato che il miglior numero di gradi di libertà è  $df = 5.93817$ . La curva è molto simile alla funzione di densità vista in precedenza. Inoltre, utilizzando gli intervalli di confidenza per calcolare la media, si è ottenuto il seguente intervallo, con un livello di confidenza pari al 95%:  $[4.613314, 6.540771]$ . La coerenza tra la stima puntuale dei gradi di libertà e l'intervallo di confidenza della media, calcolato con un livello di confidenza del 95%, fornisce ulteriore validità e supporto alla nostra analisi statistica.



Dopo aver applicato la stima puntuale anche sulla distribuzione normale, possiamo notare quanto sia diversa dalla funzione di densità mostrata in precedenza. Infatti, possiamo notare come il picco, nella distribuzione normale, si aggiri intorno al valore 5.5. Mentre nella funzione di densità è pari al 4.5 (così come nella distribuzione Chi Quadrato). Nella sezione "Test del chi quadrato", saranno utilizzate la verifica delle ipotesi e il test chi quadrato per dimostrare che la distribuzione normale non si adatta al grafico che abbiamo scelto.

## 7.3 Verifica delle ipotesi

I paesi appartenenti all'OECD (Organization for Economic Cooperation and Development) dichiarano che la percentuale di disoccupati ha una deviazione standard di 2.93. Estratto un campione di 25 paesi, si riscontra che la deviazione standard campionaria è 2.77. Si vuole verificare se la varianza dichiarata per la percentuale di disoccupazione sia effettivamente quella dichiarata. Si desidera costruire il test di misura  $\alpha = 0.05$  per verificare l'ipotesi nulla  $H_0 : \sigma^2 = 8,60$  in alternativa all'ipotesi  $H_1 : \sigma^2 \neq 8,60$ .

Occorre applicare un test di verifica di ipotesi bilaterale. Si ha che:

Livello di significatività:  $\alpha = 0.05$

Varianza dichiarata:  $\sigma_0 = 8,60$

Dimensione del campione:  $n = 25$

Deviazione standard campionaria:  $s_{25} = 2.77$

```
> alpha <- 0.05
> sigma02 <- 8.596
> n = 25
> varcamp <- 7.695
> qchisq(alpha/2, df=n-1)
[1] 12.40115
> qchisq(1-alpha/2, df=n-1)
[1] 39.36408
> (n-1)*varcamp/sigma02
[1] 21.48441
```

Si ha quindi che  $\chi^2 = 21.48$ . Poiché 21.48 cade nella regione di accettazione, occorre quindi accettare l'ipotesi nulla con un livello di significatività del 5%.

## 7.4 Test del chi quadrato

Utilizziamo il test del chi quadrato per verificare che la distribuzione normale non si adatta alla distribuzione dei dati. Innanzitutto, utilizzando i quantili della normale, vengono determinati i sottoinsiemi  $I_1, I_2, I_3, I_4$ , in modo che la probabilità che  $X$  assuma un valore appartenente a  $I_i$  ( $i = 1, 2, 3, 4$ ) sia uguale a  $p_i = 0.25$ .

Ricordando che i dati possono assumere un valore compreso tra 0 e 100, si ha che gli intervalli sono:

$$I_1 = [0, 3.53]$$

$$I_2 = (3.53, 4.44]$$

$$I_3 = (4.44, 6.79]$$

$$I_4 = (6.79, 100]$$

Le frequenze degli intervalli sono:

$$n_1 = 10, \quad n_2 = 9, \quad n_3 = 9, \quad n_4 = 10$$

Otteniamo che  $\chi^2 = 0.10526$ , e consideriamo  $\alpha = 0.05$

Dato che la regione di accettazione per il test del chi quadrato è compresa tra 0.2157953 e 9.348404, e il valore calcolato del chi quadrato rientra al di fuori di questa regione: l'ipotesi nulla, ovvero che i dati seguano una distribuzione normale, viene rifiutata.



---

## Bibliografia

---

- [1] [Online]. Available: <https://theconversation.com/its-not-just-you-work-during-covid-has-become-much-harder-150030> (Citato a pagina 4)
- [2] [Online]. Available: <https://oecdoscope.blog/2016/07/19/tackling-the-three-main-challenges-in-costa-rica-fiscal-reform-reverting-the-slowdown-in-productivity-and-reducing-inequality/> (Citato a pagina 7)
- [3] [Online]. Available: <https://savinopartners.com/2023/08/21/come-cambiata-l-economia-in-repubblica-ceca-negli-ultimi-anni/> (Citato a pagina 9)
- [4] [Online]. Available: [https://en.wikipedia.org/wiki/Greek\\_government-debt\\_crisis](https://en.wikipedia.org/wiki/Greek_government-debt_crisis) (Citato a pagina 13)
- [5] [Online]. Available: [https://it.wikipedia.org/wiki/Terremoto\\_e\\_maremoto\\_del\\_2011](https://it.wikipedia.org/wiki/Terremoto_e_maremoto_del_2011) (Citato a pagina 17)
- [6] [Online]. Available: [https://it.wikipedia.org/wiki/Incidente\\_nucleare\\_di\\_Fukushima\\_Dai-ichi](https://it.wikipedia.org/wiki/Incidente_nucleare_di_Fukushima_Dai-ichi) (Citato a pagina 17)
- [7] [Online]. Available: <https://www.economiaepolitica.it/politiche-economiche/europa-e-mondo/crisi-e-ripresa-della-lettonia-un-successo-dellausterita/> (Citato a pagina 18)

- [8] [Online]. Available: [https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms\\_424066.pdf](https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_424066.pdf) (Citato a pagina 19)
- [9] [Online]. Available: <https://www.oecd-ilibrary.org/sites/faefa993-en/index.html?itemId=/content/component/faefa993-en> (Citato a pagina 20)
- [10] [Online]. Available: <https://www.ilfattoquotidiano.it/2023/05/23/la-grecia-verso-un-nuovo-ciclo-economico-anno-record-della-borsa-ma-la-poverta-dilaga-e-disoccupati-calano-con-lavori-a-basso-reddito/7169757/> (Citato a pagina 36)
- [11] [Online]. Available: [https://st.ilsole24ore.com/rapporto-paese/lettonia/quadro-economico.shtml?refresh\\_ce=1](https://st.ilsole24ore.com/rapporto-paese/lettonia/quadro-economico.shtml?refresh_ce=1) (Citato a pagina 41)
- [12] [Online]. Available: <https://www.oecd-ilibrary.org/sites/24c13a69-en/index.html?itemId=/content/component/24c13a69-en#:~:text=Emigration%20of%20Estonian%20citizens%20to,Netherlands%20and%207%25%20to%20Germany.> (Citato a pagina 63)
- [13] [Online]. Available: <https://eige.europa.eu/gender-equality-index/2022/country/LT> (Citato a pagina 75)
- [14] [Online]. Available: <https://agente0011.it/oltre-gli-stereotipi-ruoli-di-genere/> (Citato a pagina 91)
- [15] [Online]. Available: <https://dilei.it/mamma/fare-la-casalinga-un-lavoro/468798/> (Citato a pagina 92)
- [16] [Online]. Available: [https://en.wikipedia.org/wiki/Gender\\_inequality](https://en.wikipedia.org/wiki/Gender_inequality) (Citato a pagina 92)
- [17] [Online]. Available: <https://cne.news/article/2793-czech-birth-rate-at-lowest-point-in-years> (Citato a pagina 96)
- [18] [Online]. Available: [https://en.wikipedia.org/wiki/Czech\\_diaspora](https://en.wikipedia.org/wiki/Czech_diaspora) (Citato a pagina 96)

- [19] [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/34232120/> (Citato a pagina 98)
- [20] [Online]. Available: <https://www.pwc.com/gr/en/publications/greek-thought-leadership/fertility-in-greece.html> (Citato a pagina 98)
- [21] [Online]. Available: <https://www.oecd-ilibrary.org/docserver/645ca1ac-en.pdf?expires=1710081215&id=id&accname=guest&checksum=2BE46E2BFBC4226456392AE32A889978> (Citato a pagina 99)
- [22] [Online]. Available: [https://www.quotidianosanita.it/studi-e-analisi/articolo.php?articolo\\_id=115352](https://www.quotidianosanita.it/studi-e-analisi/articolo.php?articolo_id=115352) (Citato a pagina 100)
- [23] [Online]. Available: <https://www.laregione.ch/estero/estero/1732458/figli-preoccupazioni-giappone-dichiarato-numero> (Citato a pagina 101)
- [24] [Online]. Available: <https://viaggi giappone.com/blog/giappone-lavoro.html> (Citato a pagina 101)
- [25] [Online]. Available: <https://news.bloomberglaw.com/esg/japan-urges-dads-to-make-time-for-babies-amid-population-worries> (Citato a pagina 101)
- [26] [Online]. Available: <https://www.lrt.lt/en/news-in-english/19/2088484/baby-slump-lithuania-s-fertility-rate-nearly-hits-rock-bottom> (Citato a pagina 102)
- [27] [Online]. Available: <https://www.epicentro.iss.it/ben/2012/aprile/2> (Citato a pagina 109)
- [28] [Online]. Available: [https://it.wikipedia.org/wiki/Modello\\_svedese](https://it.wikipedia.org/wiki/Modello_svedese) (Citato a pagina 131)
- [29] [Online]. Available: <https://www.iknalsemikan.com/post/work-life-balance-discover-why-sweden-is-a-great-place-to-work> (Citato a pagina 132)
- [30] [Online]. Available: <https://en.wikipedia.org/wiki/Flextime> (Citato a pagina 132)

- [31] [Online]. Available: <https://www.psicologiacontemporanea.it/blog/i-giovani-oggi-non-hanno-voglia-vogliono-lavorare-ma-e-davvero-cosi/> (Citato a pagina 132)
- [32] [Online]. Available: <https://www.oecd-ilibrary.org/sites/9789264269231-10-en/index.html?itemId=/content/component/9789264269231-10-en#:~:text=According%20to%20the%20Central%20Bank,economic%20activity%20and%20institutional%20sector.> (Citato a pagina 133)
- [33] [Online]. Available: [https://it.wikipedia.org/wiki/Fuga\\_di\\_cervelli](https://it.wikipedia.org/wiki/Fuga_di_cervelli) (Citato a pagina 134)
- [34] [Online]. Available: <https://www.avvisopubblico.it/home/home/cosa-facciamo/informare/osservatorio-parlamentare/attivita-dinchiesta/lattivita-di-inchiesta-parlamentare-nella-xviii-legislatura/relazione-sulle-condizioni-di-lavoro-in-italia-sullo-sfruttamento-e-sulla-sicurezza-nei-luoghi-di-lavoro-pubblici-e-privati-scheda-di-sintesi/> (Citato a pagina 134)