



ELABORATO DI STATISTICA SOCIALE 2022/2023

Anderloni Hanan Francesca 889079

Caruso Giusy 886460

Finzi Rebecca 882523

Mato Ilaria 882334

ESERCIZIO 1

L'ingrediente principale di una ricerca è lo schema concettuale, da cui viene tradotto il questionario. L'organizzazione delle informazioni contenute nello schema avviene mediante "albero delle aree": rappresentazione di fenomeni o entità pertinenti alla ricerca, creando legami logici e rappresentando quelle che sono le variabili delle entità.

Facendo riferimento alla bozza "Questionario Sport" [1] le aree che sono emerse attraverso le domande sono:

1. Opinioni (domande numero 9, 13, 14, 15, 23, 27, 29, 30, 31, 32)
2. Intenzioni comportamentali (domande numero 17, 18, 19, 22, 26, 33)
3. Stati d'animo (domande numero 8, 24, 28)
4. Comportamenti fattuali (domande numero 6, 7, 10, 11, 12, 16, 21, 25, 34)
5. Caratteristiche socio-demografiche (domande numero 1, 2, 3, 4, 5)

Le aree sono diverse tra loro e hanno una propria interpretazione: l'albero delle aree (Grafico 1) oltre ad essere visibilmente comprensibile permette di identificare un percorso ottimale tra esse, basandosi sul principio dell'indipendenza della compilazione, ovvero ogni area è compilata indipendentemente dalla compilazione delle altre.

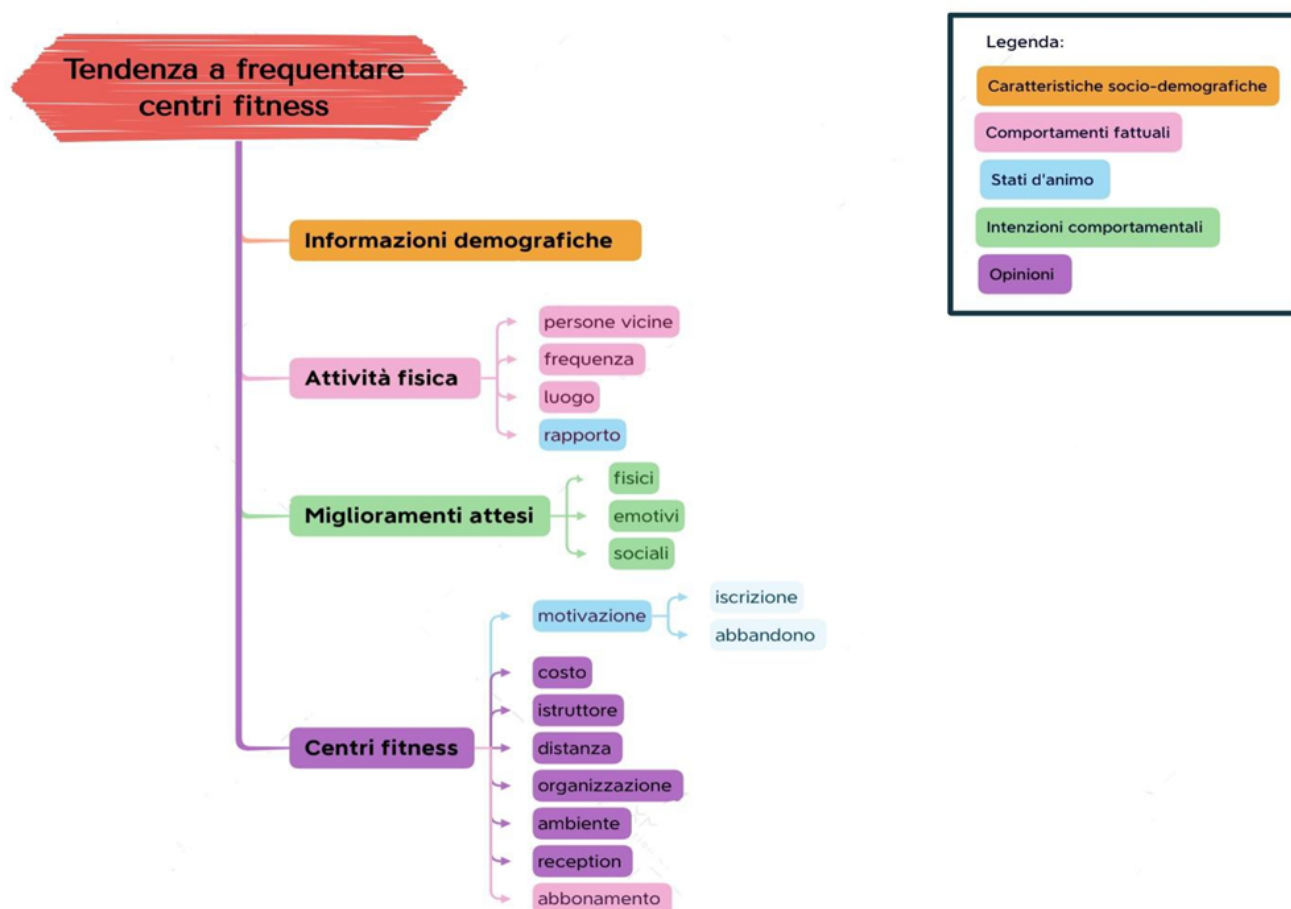


Grafico 1 Albero delle aree. Fonte: Questionario Sport.

All'inizio del questionario vengono poste delle domande conoscitive che permettono di avere informazioni generali sul rispondente e, in linea generale, il suo rapporto con l'attività fisica. Le domande seguenti trattano i miglioramenti attesi dal punto di vista fisico, sociale o mentale rispetto all'attività praticata.

Successivamente il focus si sposta sull'attività che si può svolgere all'interno di centri sportivi/club, sui quali si chiede l'opinione dell'intervistato riguardo il costo, il personale e l'ambiente delle strutture.

L'obiettivo della ricerca è dunque analizzare e capire la predisposizione degli intervistati ad iscriversi o meno ad un centro fitness, valutando tutti gli aspetti sia tecnici che personali.

L'area analizzata delle categorie di domande del questionario Sport è relativa alle intenzioni comportamentali dei soggetti interessati, di cui fanno parte le domande riportate in seguito.

- **17.** Quali sono le Vostre aspettative per un miglioramento delle condizioni fisiche?.....
.....
- **18.** Quali sono le Vostre aspettative per un miglioramento delle condizioni di equilibrio psicofisico(stress)?.....
.....
- **19.** Quali sono le Vostre aspettative per un miglioramento delle relazioni sociali?.....
.....

Queste domande si riferiscono ad un'intenzione comportamentale nel futuro. Si chiede al rispondente di esprimersi riguardante ad una tematica personale come la condizione psicologica, fisica e sociale che sono questioni di cui si può risentire particolarmente, quindi spesso possono essere difficile da affrontare.

Il suggerimento sarebbe quello di unire insieme queste domande in un unico blocco e porle in modo tale da poter rispondere in modo chiuso. Così le risposte sarebbero predefinite, per esempio utilizzando il metodo delle scale di giudizio: accordo/disaccordo, gradimento (positivo o negativo), valutazione in termini di grandezza percentuale e frequenza, si semplifica la risposta mettendo a proprio agio l'intervistato. Cambiando il metodo di risposta andranno anche riformulate le domande.

- **22.** Per frequentare un centro fitness o centro sportivo e soddisfare i Vostri bisogni siete disposti a spendere:
al mese €..... Ogni tre mesi €..... Ogni sei mesi €..... all'anno €.....

Anche la domanda 22 risulta aperta, ma bisogna sottolineare degli errori da evitare. Poiché la domanda per il suo contenuto risulta essere intrusiva, invece di chiedere il costo preciso per l'abbonamento che il rispondente sarebbe disposto a spendere, sarebbe meglio mettere delle fasce di costo. Così invece ci si limita ad una precisazione non necessaria, che può essere

evitata non mettendo sotto sforzo l'interlocutore. Quindi per poter ovviare a questa problematica basterebbe mettere una scala ordinale gerarchica.

Una seconda problematica potrebbe essere la terminologia utilizzata: "centro sportivo/centro fitness". Fin dall'inizio del questionario è stata chiamata "palestra" o "club" ma cambiare la terminologia a metà dell'intervista può creare complicazioni perché si perde l'omogeneità nel linguaggio. Sarebbe meglio avvertire all'inizio del questionario di eventuali cambiamenti nel percorso, per non confondere.

- **26.** Quale mezzo usate per raggiungere il centro fitness?..... e quale distanza siete disposti a percorrere?.....

La domanda non può essere definita singola, perché al suo interno presenta due domande, rendendola una domanda doppia. La difficoltà in questo caso è che il rispondente può rispondere in maniera diversa nella prima e/o nella seconda domanda, ma se sono nella stessa domanda, non si possono dare risposte e opportuni accorgimenti per ognuna di loro.

Il suggerimento sarebbe quello di separarle e renderle più generiche. Perché chiedendo in modo aperto con quale mezzo raggiungo il centro fitness, potrebbe esserci una lunga lista (che siano tutti i mezzi personali oppure i mezzi di trasporto pubblici). Quindi si potrebbe mettere una lista generica di scelte multiple, tra le quali si possono scegliere più mezzi, senza lasciare al rispondente l'impegno di scrivere tutti i possibili mezzi che potrebbe usare.

Inoltre la seconda domanda, che andrebbe separata dalla prima, è sbagliata, infatti sarebbe meglio mettere un range di valori su una scala, per darne solo un'idea generale della distanza e non una precisazione inutile.

- **33.** Quali sono le Vs. aspettative all'inizio dell'anno?

Domanda aperta troppo generica rispetto all'obiettivo della ricerca. Lascia troppa libertà di risposta all'intervistato chiedendogli di reperire informazioni poco chiare che avrà difficoltà ad esprimere. Bisognerebbe mettere una precisazione dell'argomento nella domanda, perché non è chiaro a quale aspettativa fa riferimento, come: il fisico, la competenza dei personal trainer, la struttura della palestra.

Inoltre se si decide di lasciarla aperta, si va incontro al rischio di dilungarsi troppo nella risposta, rischiando di dirigere l'intervista verso altre tematiche non richieste con enormi digressioni, allungando notevolmente il tempo del colloquio.

ESERCIZIO 2

punto 2.1

Il calcolo dei pesi si trova nell'Excel "Anderloni_Caruso_Finzi_Mato" nel foglio "pesi_es2".

punto 2.2

Utilizzando i dati presenti nel dataset "domandeverie 2018-22" [2] sono state costruite, con il programma SPSS, due scale dell'aggressività con il metodo di Thurstone.

Per creare le seguenti scale abbiamo considerato un certo numero di items, a seconda del nostro scopo.

Per la prima scala, chiamata "scala_aggressività_1", abbiamo utilizzato 15 items con i loro relativi pesi. In particolare, sono stati considerati gli items: V18_1, V18_8, V18_17, V18_22, V18_29, V18_33, V18_37, V18_38, V18_40, V18_43, V18_44, V18_54, V18_55, V18_59, V18_60.

Invece per la seconda scala, chiamata "scala_dell'aggressività_2", abbiamo considerato solo 7 items, tra i 15 sopra elencati, con i relativi pesi. In particolare, sono stati selezionati sono gli items strettamente legati all'aggressività esterna, cioè gli items: V18_1, V18_8, V18_22, V18_43, V18_54, V18_55, V18_60. Quindi la seconda rispetto alla prima scala misura l'aggressività esterna mentre la prima è più generale, e misura sia quella esterna che interna.

In primo luogo, abbiamo studiato le distribuzioni e le principali caratteristiche delle due scale come valori di sintesi, percentili e simmetria.

Statistics		
scala_aggressività_1		
N	Valid	461
	Missing	90
Mean		9.5046
Median		9.4521
Mode		11.39
Std. Deviation		1.15263
Variance		1.329
Range		6.49
Minimum		6.49
Maximum		12.97
Percentiles	25	8.7046
	50	9.4521
	75	10.2686

Tabella 1 Statistiche descrittive della scala dell'aggressività 1. Fonte: dataset "domandeverie 2018-22".

Statistics		
scala_aggressività_2		
N	Valid	551
	Missing	0
Mean		7.0543
Median		6.8895
Mode		6.35
Std. Deviation		.96950
Variance		.940
Range		4.59
Minimum		5.12
Maximum		9.71
Percentiles	25	6.3512
	50	6.8895
	75	7.7052

Tabella 2 Statistiche descrittive della scala dell'aggressività 2. Fonte: dataset "domandeverie 2018-22".

Dalla analisi descrittiva riportata si evince che nella scala aggressività 1 (Tabella 1) si possono notare 90 valori mancanti (*missing data*), cioè il 16% dei dati sono mancanti, alcune persone dunque si sono astenute dal rispondere a certe domande.

Questo non accade invece nella scala dell'aggressività 2 (Tabella 2) dove i *missing data* risultano pari a zero, ovvero tutti i rispondenti si sono impegnati a rispondere a tutte le domande. Infatti, consultando le risposte del questionario in *data view* su SPSS si può notare che all'item V18_29, considerato per la costruzione della scala 1 e non per quella della scala 2, molte persone non hanno risposto.

Nella prima scala dell'aggressività (Tabella 1) si osserva un valore medio più alto, di 9.5046, rispetto al valore medio della seconda scala (Tabella 2) pari a 7.0543. La prima scala presenta inoltre una varianza maggiore rispetto alla seconda, rispettivamente 1.329 della prima contro 0.940 della seconda. Questo significa che la dispersione attorno al valore medio della prima scala è superiore rispetto a quella della scala 2. Infatti, tanto maggiore è la varianza tanto più i valori sono dispersi attorno al valore medio, invece tanto minore è la varianza tanto più i valori della distribuzione sono concentrati attorno al valore medio.

Un'altra quantità calcolata nell'analisi che misura la dispersione è il range, però a differenza della varianza non ci dice se i dati sono concentrati attorno a un certo valore medio ma solamente quanto sono dispersi. Infatti, è una misura che non indica come, ma solo quanto sono dispersi i dati. In particolare, il range indica l'intervallo di variazione della variabile considerata, ed è dato dalla differenza tra il valore massimo e minimo, siccome si basa solo su due osservazioni è fortemente influenzato dai valori estremi. Nella scala dell'aggressività 1 il range risulta 6.49, mentre nella scala 2 risulta inferiore, con un valore di 4.59, quindi nella prima scala i dati sono più dispersi rispetto alla seconda.

Studiando assieme il valore della media, moda e mediana si può determinare se la scala sotto analisi gode di una situazione di simmetria, asimmetria positiva o asimmetria negativa. Dai risultati descrittivi si può vedere che nel caso della scala dell'aggressività 1 (Tabella 1) c'è una situazione di leggera asimmetria negativa, determinata non tanto dalla media e dalla moda che si aggirano entrambe allo stesso valore di 9.50 ma per la moda, la quale risulta essere molto più grande con un valore di 11.39. Invece nel caso della scala dell'aggressività esterna (Tabella 2) vediamo la distribuzione tendente a un'asimmetria positiva in quanto la media (7.05) è maggiore della mediana (6.9) che è a sua volta maggiore della moda (6.4).

Dopodiché un'altra informazione che ci fornisce questa analisi descrittiva fatta su SPSS sono informazioni relative ai quartili, valori che si ottengono dividendo l'insieme di dati ordinati in 4 parti uguali. Il 1° quartile o 25° percentile è il valore che lascia alla sua sinistra il 25% degli elementi della distribuzione. Il secondo quartile o 50° percentile è il più importante poiché coincide con la mediana; infatti, è il valore che lascia alla sua sinistra il 50% della distribuzione. Il terzo quartile o 75° percentile è il valore che lascia il 75% degli elementi a sinistra e il 25% a destra.

Nella scala dell'aggressività 1 (tabella 1) il primo quartile è pari a 8.70, cioè almeno un quarto della popolazione ha un'aggressività inferiore o uguale a 8.70, il secondo quartile o mediana è 9.50 quindi almeno la metà della popolazione presenta un valore minore o uguale a 9.50, infine il terzo quartile è 10.27. Mentre nel caso della scala dell'aggressività esterna, il primo quartile è 6.35, il secondo è 6.89, il terzo è 7.70.

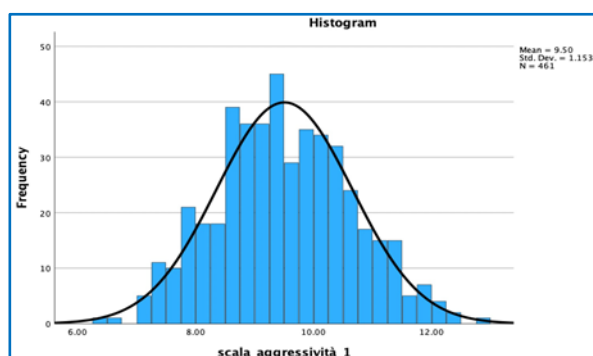


Grafico 1 distribuzione della scala dell'aggressività 1 Fonte: dataset "domandeverie 2018-22".

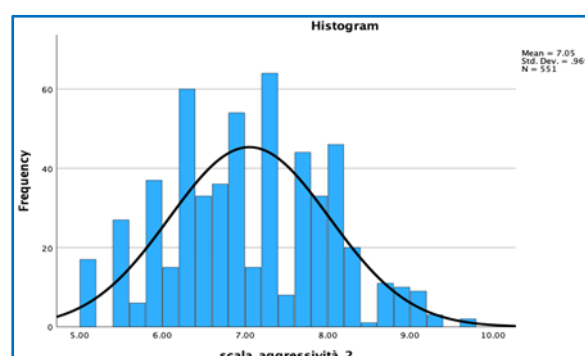


Grafico 2 distribuzione della scala dell'aggressività 2. Fonte: dataset "domandeverie 2018-22".

Dal punto di vista grafico confrontando la distribuzione delle due scale rispetto una distribuzione normale si può notare che gli istogrammi delle due scale di aggressività sono distribuiti in modo diverso. Si vede chiaramente che il grafico della scala dell'aggressività 1 (Grafico 1) segue maggiormente la distribuzione Normale, rispetto al grafico della scala dell'aggressività 2 (Grafico 2) in cui le frequenze dei valori sono maggiormente discontinue e oscillanti. Una delle conclusioni che si può trarre è che la scala dell'aggressività 1 ha dati più coerenti con la distribuzione Normale, nonostante abbia una deviazione standard maggiore rispetto al secondo istogramma.

punto 2.3

2.3.1 confronto maschi e femmine

Segue il confronto dell'aggressività a seconda del sesso femminile o maschile.

scala aggressività 1

Case Processing Summary						
	Included		Cases Excluded		Total	
	N	Percent	N	Percent	N	Percent
scala_aggressività_1 *	461	83.7%	90	16.3%	551	100.0%
Sesso						

Report			
scala_aggressività_1			
Sesso	Mean	N	Std. Deviation
maschio	9.5150	237	1.18125
femmina	9.4936	224	1.12408
Total	9.5046	461	1.15263

Tabella 3 Confronto aggressività a seconda del sesso nella scala dell'aggressività 1. Fonte: dataset "domandeverie 2018-22".

Dalle analisi risulta che le donne, anche se di poco, siano meno aggressive degli uomini, rispettivamente con una media di 9.49 contro una di 9.52 (Tabella 3).

Per capire se la differenza è statisticamente significativa o meno, facciamo un test per confrontare le medie, come il test t o il test Anova, a seconda del numero di popolazioni.

In questo caso si preferisce fare il test t e non il test ANOVA perché il test t è uno strumento per valutare se c'è una differenza statisticamente significativa tra le medie di due gruppi tra loro indipendenti e in questo caso abbiamo solo due popolazioni: cioè la media dell'aggressività delle donne e quella dei maschi.

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Significance One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower Upper
scala_aggressività_1	Equal variances assumed	.266	.606	.198	459	.421	.843	.02131	.10752	-.18998 .23261
	Equal variances not assumed			.198	458.978	.421	.843	.02131	.10737	-.18969 .23231

Tabella 4 Test t per il confronto delle medie nella scala dell'aggressività 1 a seconda del sesso. Fonte: dataset "domandeverie 2018-22".

Dal test t (Tabella 4) si può vedere che le due popolazioni, cioè maschi e femmine, hanno la stessa media perché il valore zero è compreso nell'intervallo di confidenza; quindi, concludiamo che la differenza osservata non è statisticamente significativa.

Si può notare che anche il test F, per il confronto tra due varianze, conferma l'analisi precedente: la significatività (Sig.) maggiore di 0.05 afferma che le due popolazioni derivano da una stessa popolazione con uguale media parametrica.

scala aggressività 2

Case Processing Summary						
	Included		Cases Excluded		Total	
	N	Percent	N	Percent	N	Percent
scala_aggressività_2 *	551	100.0%	0	0.0%	551	100.0%
Sesso						

Report			
scala_aggressività_2			
Sesso	Mean	N	Std. Deviation
maschio	6.9843	277	1.02908
femmina	7.1250	274	.90166
Total	7.0543	551	.96950

Tabella 5 Confronto aggressività a seconda del sesso nella scala dell'aggressività 2. Fonte: dataset "domandevarie 2018-22".

Utilizzando la scala dell'aggressività esterna (scala aggressività 2) si osserva che i risultati si invertono rispetto alla scala dell'aggressività 1 e che le donne sembrano essere leggermente più aggressive degli uomini, con rispettivamente una media di 6.9843 e di 7.1250 (Tabella 5). Anche per la scala dell'aggressività esterna guardiamo se tale differenza è statisticamente significativa con il test t di Student.

Independent Samples Test											
		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
scala_aggressività_2	Equal variances assumed	3.087	.079	-1.706	549	.044	.089	-.14068	.08246	-.30266	.02130
	Equal variances not assumed			-1.707	541.129	.044	.088	-.14068	.08240	-.30255	.02119

Tabella 6 Test t per il confronto delle medie nella scala dell'aggressività 2 a seconda del sesso. Fonte: dataset "domandevarie 2018-22".

Grazie al test t (Tabella 6) possiamo affermare che le medie della popolazione delle femmine e della popolazione dei maschi sono uguali poiché lo zero è compreso nell'intervallo di confidenza.

2.3.2 confronto segni zodiacali

scala aggressività 1

Case Processing Summary						
	Included		Cases Excluded		Total	
	N	Percent	N	Percent	N	Percent
scala_aggressività_1 * Segno zodiacale	461	83.7%	90	16.3%	551	100.0%

Report			
scala_aggressività_1			
Segno zodiacale	Mean	N	Std. Deviation
ariete	9.7137	36	1.13826
toro	9.3869	31	1.34134
gemelli	9.6526	43	1.12402
cancro	9.4073	40	1.21566
leone	9.3784	40	1.06622
vergine	10.0364	40	1.22962
bilancia	9.2530	38	1.34128
scorpione	9.4002	47	.87527
sagittario	9.2501	39	1.01838
capricorno	9.6074	34	1.06794
acquario	9.4717	36	1.20050
pesci	9.4912	37	1.14737
Total	9.5046	461	1.15263

Tabella 7 Confronto aggressività a seconda del segno zodiacale nella scala dell'aggressività 1. Fonte: dataset "domandevarie 2018-22".

Analizzando l'aggressività nei segni zodiacali il segno più aggressivo risulta essere quello della vergine con una media del 10.04 mentre quello meno aggressivo, con una media del 9.25, risulta essere quello del sagittario (Tabella 7).

Per capire se la differenza è statisticamente significativa o meno in questo caso, dove le popolazioni non sono più due ma dodici abbiamo preferito fare il test ANOVA perché quest'ultimo è un metodo statistico per testare le differenze tra le medie di tre o più gruppi.

ANOVA					
scala_aggressività_1					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21.122	11	1.920	1.461	.143
Within Groups	590.009	449	1.314		
Total	611.131	460			

Tabella 8 Test ANOVA per il confronto delle medie nella scala dell'aggressività 1 a seconda del segno zodiacale. Fonte: dataset "domandevarie 2018-22".

Dal test ANOVA (Tabella 8) risulta che le medie non hanno differenze significative, dato che la significatività è pari a 0.143 e quindi maggiore di 0.05.

scala aggressività 2

Case Processing Summary						
	Included		Cases Excluded		Total	
	N	Percent	N	Percent	N	Percent
scala_aggressività_2 * Segno zodiacale	551	100.0%	0	0.0%	551	100.0%

Report			
scala_aggressività_2			
Segno zodiacale	Mean	N	Std. Deviation
ariete	7.1754	45	.93638
toro	7.0372	37	1.05438
gemelli	7.1123	53	.94651
cancro	7.0153	46	1.03953
leone	7.0089	47	.92850
vergine	7.3786	49	1.08551
bilancia	6.8578	47	1.12796
scorpione	7.0719	57	.85031
sagittario	6.8757	45	.70573
capricorno	7.0128	40	.80306
acquario	7.0790	42	1.06739
pesci	6.9850	43	1.02808
Total	7.0543	551	.96950

Tabella 9 Confronto aggressività a seconda del segno zodiacale nella scala dell'aggressività 2. Fonte: dataset "domandevarie 2018-22".

Nella scala dell'aggressività esterna (Tabella 9) il segno della vergine è sempre quello più aggressivo, con una media di 7.3786, mentre quello meno aggressivo è quello della bilancia con una media di 6.8578. Anche in questo caso utilizziamo il test ANOVA per vedere se questa differenza di aggressività tra i segni zodiacali è statisticamente significativa o è dovuta al caso.

ANOVA					
scala_aggressività_2					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	9.740	11	.885	.941	.500
Within Groups	507.219	539	.941		
Total	516.959	550			

Tabella 10 Test ANOVA per il confronto delle medie nella scala dell'aggressività 1 a seconda del segno zodiacale. Fonte: dataset "domandevarie 2018-22".

Dal test (Tabella 10) risulta che le medie non sono significativamente diverse perchè la significatività è maggiore di 0.05 in particolare è pari a 0.5, quindi la differenza osservata tra le medie dei segni zodiacali è dovuta solamente al caso.

2.3.3 Confronto tra ottimismo e scala aggressività

Abbiamo creato una nuova variabile denominata "scala_ottimismo_finale" (Tabella 11) in modo additivo, ovvero andando a sommare tutte 6 le variabili di interesse.

I dati tratti dalle risposte alle domande del questionario “domandevarie 2018-22”, sono basati sulla risposta ad una scala nominale che va da 1 a 5, dove il massimo del punteggio rappresenta l’ottimismo, mentre un punteggio più basso definisce l’individuo pessimista (1=pessimista, 2=un po’ pessimista, 3=né pessimista né ottimista, 4=un po’ ottimista, 5=ottimista). Si è prestato attenzione a correggere eventuali items in cui, al contrario, più alto era il punteggio più un soggetto era pessimista, e in tal caso è stato invertito il verso con una ricodifica. In particolare, sono state considerate tre variabili di partenza e tre ricodificate. Così una volta ottenuti tutti gli items nello stesso verso, l’analisi si è potuta svolgere in modo più semplice e coerente.

Statistics		
scala_ottimismo_finale		
N	Valid	551
	Missing	0
Mean		17.9129
Median		18.0000
Mode		16.00
Std. Deviation		4.43208
Variance		19.643
Range		24.00
Minimum		6.00
Maximum		30.00
Percentiles	25	15.0000
	50	18.0000
	75	21.0000

Tabella 11 Statistiche descrittive della scala_ottimismo_finale.

Fonte: dataset “domande varie 2018-22”

Scala aggressività 1

Per fare un confronto tra le due scale dell’aggressività 1 e la scala dell’ottimismo è stato fatto in primo luogo un confronto tra i rispettivi istogrammi, di tipo grafico.

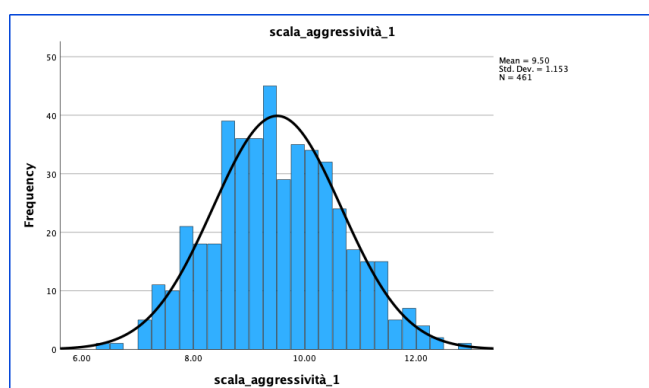


Grafico 3 Distribuzione della scala dell’aggressività 1. Fonte: dataset “domandevarie 2018-22”.

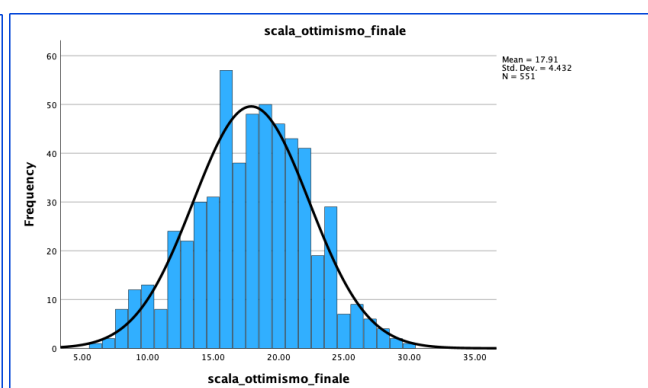


Grafico 4 Distribuzione della scala dell’ottimismo. Fonte: dataset “domandevarie 2018-22”.

La media relativa alla scala dell’aggressività 1 è di 9.50 mentre quella relativa alla scala dell’ottimismo è pari a 17.91, inoltre la scala dell’aggressività 1 si adatta meglio alla distribuzione Normale rispetto alla scala dell’ottimismo (Grafici 3 e 4). Le deviazioni standard

della scala di aggressività e quella dell'ottimismo sono rispettivamente di 1.153 e 4.432. La deviazione standard della scala dell'ottimismo è maggiore rispetto a quella dell'aggressività poiché i valori attorno alla media sono più grandi. Infatti, la deviazione standard è un indice che indica quanto i numeri si distanziano dalla media aritmetica. Quindi dove, come nel caso della scala dell'ottimismo, i valori sono più lontani dalla media, sarà più grande la deviazione standard.

Per vedere se la differenza osservata tra la scala dell'aggressività e la scala dell'ottimismo è statisticamente significativa, è stato utilizzato il test del Chi-Quadrato che stabilisce se le due variabili sono indipendenti.

Chi-Square Tests			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	9074.424 ^a	8901	.097
Likelihood Ratio	2462.164	8901	1.000
Linear-by-Linear Association	22.673	1	<.001
N of Valid Cases	461		
a. 9312 cells (100.0%) have expected count less than 5. The minimum expected count is .00.			

Tabella 12 Test Chi-Quadrato per l'indipendenza tra la scala dell'aggressività 1 e la scala dell'ottimismo. Fonte: dataset "domandeverie 2018-22".

Dalla tabella 12 vediamo che il valore della voce *Asymptotic Significance*, 0.097, è maggiore di 0.05 quindi si può affermare che la differenza tra le due scale non è statisticamente significativa. Infatti, poiché il p-value, cioè 0.097, è maggiore di 0.05 non posso rifiutare l'ipotesi nulla (Ho) di indipendenza tra le variabili.

Dopodiché è stata calcolata anche la correlazione lineare per vedere quanto la relazione tra le due variabili è lineare.

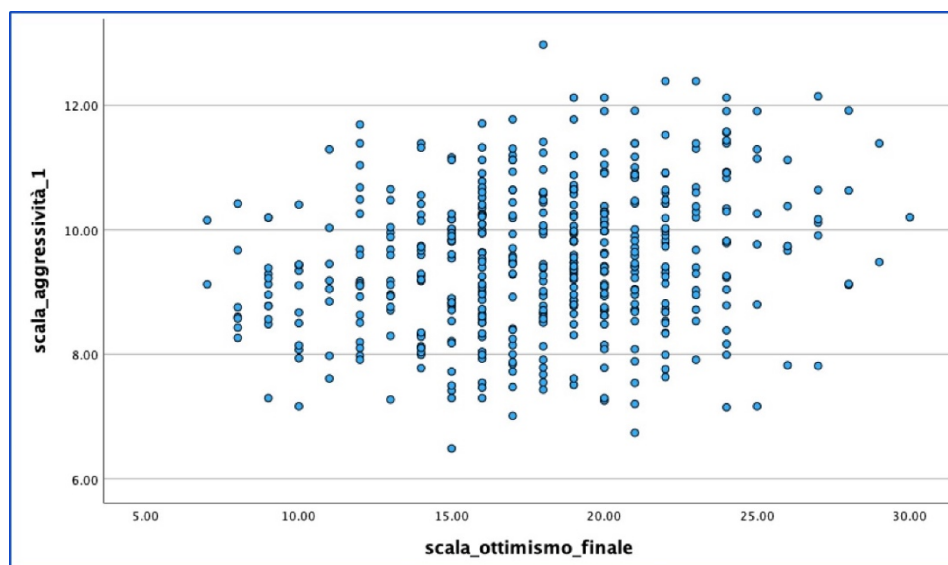
Correlations			
		scala_aggressività_1	scala_ottimismo_finale
scala_aggressività_1	Pearson Correlation	1	.222**
	Sig. (2-tailed)		<.001
	N	461	461
scala_ottimismo_finale	Pearson Correlation	.222**	1
	Sig. (2-tailed)	<.001	
	N	461	551
**. Correlation is significant at the 0.01 level (2-tailed).			

Tabella 13 Correlazione tra scala dell'aggressività 1 e scala dell'ottimismo. Fonte: dataset "domandeverie 2018-22".

Poiché il valore dell'indice *Pearson Correlation* (Tabella 13) tra la scala dell'aggressività 1 e la scala dell'ottimismo è di 0.22, possiamo affermare che la correlazione è molto bassa. Nonostante ci sia un legame lineare molto basso e quasi trascurabile, il valore è positivo

poiché compreso tra 0 e 1, quindi al crescere dell'ottimismo aumenta leggermente l'aggressività.

L'analisi grafica seguente attraverso il grafico a dispersione (Grafico 5), conferma gli esiti raggiunti finora. Infatti, la nuvola dei punti rende plausibile che la correlazione presenti un valore molto basso e l'elevata dispersione dei punti conferma l'inesistenza di una relazione lineare. Nonostante ciò, possiamo notare un leggero andamento positivo dei dati.



*Grafico 5 Grafico a dispersione dell'aggressività al variare dell'ottimismo.
Fonte: dataset "domandeverie 2018-22".*

scala aggressività 2

Lo stesso ragionamento è stato fatto anche per il confronto con la scala dell'aggressività esterna e la scala dell'ottimismo. Anche in questo caso l'analisi è stata iniziata con un confronto di tipo grafico (Grafici 6 e 7).

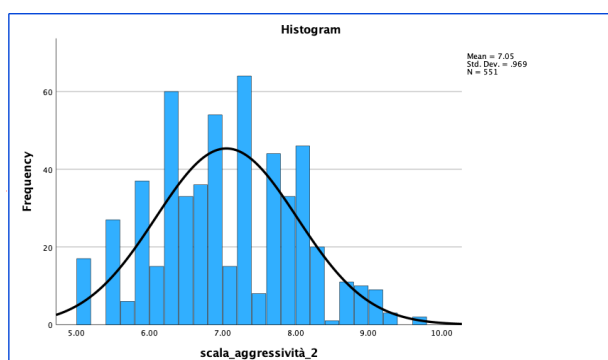


Grafico 6 Distribuzione della scala dell'aggressività 2. Fonte: dataset "domandeverie 2018-22".

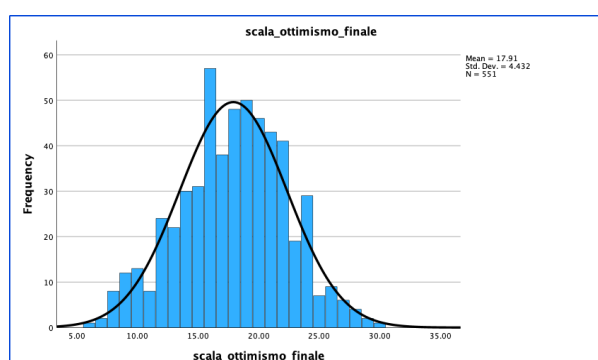


Grafico 7 Distribuzione della scala dell'ottimismo. Fonte: dataset "domandeverie 2018-22".

La media relativa alla scala dell'aggressività è di 7.05 mentre quella relativa alla scala dell'ottimismo è di 17.91, in questo caso al contrario dell'analisi precedente è la scala dell'ottimismo che si adatta meglio alla distribuzione Normale rispetto alla scala

dell'aggressività esterna. La deviazione standard della scala di aggressività esterna e della scala dell'ottimismo è rispettivamente di 0.969 e 4.432. La grande differenza tra questi due valori è data dal fatto che gli scarti dalla media nella scala dell'aggressività esterna sono minori rispetto agli scarti dalla media nella scala dell'ottimismo. Si può notare infine che, in questo caso, il numero 'N' di rispondenti è in entrambe le scale 551 poiché non abbiamo nessun valore mancante.

Chi-Square Tests			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1759.928 ^a	1512	<.001
Likelihood Ratio	921.590	1512	1.000
Linear-by-Linear Association	8.102	1	.004
N of Valid Cases	551		
a. 1599 cells (99.9%) have expected count less than 5. The minimum expected count is .00.			

Tabella 14 Test Chi-Quadrato per l'indipendenza tra la scala dell'aggressività 2 e la scala dell'ottimismo. Fonte: dataset "domandeverie 2018-22".

Poiché il valore di *Asymptotic Significance* (Tabella 14), cioè il p-value, è un valore minore di 0.05 si è portati a rifiutare l'ipotesi nulla di indipendenza e si accetta l'ipotesi alternativa di dipendenza. La differenza tra le due scale risulta essere statisticamente significativa.

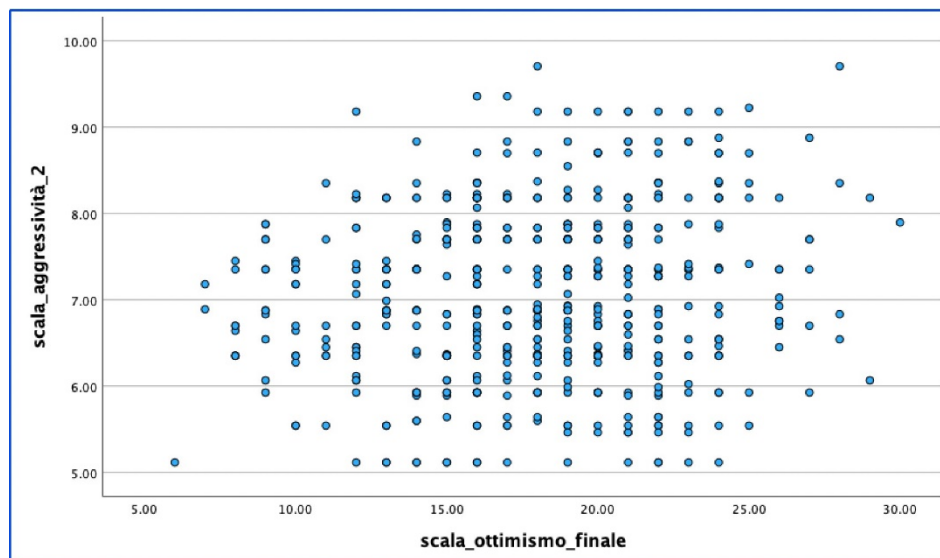
Correlations			
		scala_aggressività_2	scala_ottimismo_finale
scala_aggressività_2	Pearson Correlation	1	.121**
	Sig. (2-tailed)		.004
	N	551	551
scala_ottimismo_finale	Pearson Correlation	.121**	1
	Sig. (2-tailed)	.004	
	N	551	551
**. Correlation is significant at the 0.01 level (2-tailed).			

Tabella 15 Correlazione tra la scala dell'aggressività 2 e la scala dell'ottimismo. Fonte: dataset "domandeverie 2018-22".

Per quanto riguarda la scala aggressività esterna (scala dell'aggressività 2), è stato ottenuto un valore della correlazione di 0.121 (Tabella 15), ulteriormente inferiore a quello della scala dell'aggressività generale (scala dell'aggressività 1). Di conseguenza anche in questo caso il legame lineare è assente. Si ipotizza la presenza di un altro tipo di legame tra due variabili. Tuttavia la correlazione lineare, sebbene sia bassa, è compresa tra 0 e 1, quindi al crescere l'ottimismo cresce leggermente l'aggressività.

La rappresentazione del grafico a nuvola (Grafico 8) è coerente agli esiti dell'analisi, si conferma nuovamente la bassa correlazione: c'è una grande dispersione dei punti che non

segue l'andamento di una retta. Sebbene la retta non sia adatta a descrivere la nuvola di punti, c'è comunque un piccolo e quasi inesistente accenno di andamento lineare.



*Grafico 8 Grafico a dispersione dell'aggressività al variare dell'ottimismo.
Fonte: dataset "domandevarie 2018-22".*

ESERCIZIO 3

Punto a

La domanda 12 (variabile V13) era la seguente:

“Sara ha 32 anni, è single e vive da sola. Ha una laurea in filosofia e durante l'università era molto sensibile ai temi della giustizia sociale e partecipava a manifestazioni contro il nucleare. Secondo te è più probabile:

1. Sara è impiegata di banca
2. Sara è impiegata di banca e ed è attiva nel movimento per le pari opportunità”

Abbiamo analizzato le risposte a questa domanda in modo da verificare se ci fosse stato o meno un errore sistematico di attribuzione. Infatti, in domande di questo tipo è facile cadere nell'euristica della rappresentatività, la quale fa sì che ci venga naturale scegliere la seconda risposta. L'euristica della rappresentatività è una scorciatoia umana che in alcune situazioni può esserci utile ma che in certi casi fa incorrere in un errore.

In questo caso quando viene chiesto di scegliere tra le due opzioni, invece di pensare in maniera razionale e rispondere che l'evento più probabile è il primo e non il secondo, molti scelgono la seconda opzione, che però essendo l'intersezione di due eventi ha meno probabilità di accadere.

1. Abbiamo prima di tutto guardato in base alle risposte delle persone se avessero utilizzato o meno l'euristica della rappresentatività.

Sara/euristica rappr					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	banca+pari opport	435	78.9	78.9	78.9
	banca	116	21.1	21.1	100.0
	Total	551	100.0	100.0	

Tabella 16 Frequenza delle risposte alla domanda 12 (V13).

Fonte: dataset “domandevarie 2018-22”.

Come si può notare dalla tabella 16 probabilmente è stata usata l'euristica della rappresentatività poiché guardando le percentuali risulta che il 78% della gente che ha compilato il questionario ha risposto con 'banca + pari opportunità' e non con la risposta giusta cioè solamente 'banca'.

2. Successivamente siamo andati a vedere che differenza emergesse tra gli studenti di statistica e non, per vedere se gli studenti di statistica avessero ragionato in modo più razionale rispetto agli altri.

Sara/euristica rappr					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	banca+pari opport	317	82.1	82.1	82.1
	banca	69	17.9	17.9	100.0
	Total	386	100.0	100.0	

*Tabella 17 Frequenza delle risposte degli studenti di statistica alla domanda 12 (V13).
Fonte: dataset "domandevarie 2018-22".*

Così abbiamo fatto la stessa analisi utilizzando solo le risposte degli studenti di statistica (Tabella 17), dove risulta che la percentuale di persone che ha usato la euristica della rappresentatività sono l'82%. Quindi la frequenza dell'uso dell'euristica nel ragionamento risulta essere indipendente da chi risponde.

- Infine, abbiamo guardato se tra gli studenti di statistica ci fosse una differenza tra i diversi anni di corso (*wave*) cioè se ci fosse stata un'annata in cui si è stati più o meno razionali.

wave * Sara/euristica rappr Crosstabulation				
		Sara/euristica rappr		
Count		banca+pari opport	banca	Total
wave	1.00	68	16	84
	2.00	56	11	67
	3.00	45	14	59
	4.00	64	12	76
	5.00	84	16	100
Total		317	69	386

Tabella 18 Frequenza delle risposte alla domanda 12 (V13) per anno (dal 2019/2020 al 2020/2021). Fonte: dataset "domandevarie 2018-22".

Abbiamo ottenuto gli stessi risultati per l'anno 2019/2020, 2021/2022 e 2022/2023 in cui solo il 16% degli studenti ha risposto correttamente, mentre l'84% ha risposto in maniera sbagliata per colpa dell'euristica. Un po' meglio è andato il primo anno in esame, cioè il 2018/2019 in cui il 19% degli studenti ha risposto correttamente a fronte di un 81% che ha risposto in modo sbagliato. Mentre gli studenti che hanno risposto meglio a questa domanda sono quelli dell'anno scolastico 2020/2021: il 24% ha risposto correttamente e il 76% ha risposto sbagliato (Tabella 18).

Punto b

Nel punto b abbiamo analizzato le variabili V15, V16_2 e V17 per indicare se è plausibile dire che il campione ha risposto alla domanda 16 (variabile V17) facendo ricorso all'euristica dell'ancoraggio.

Nel questionario è stato chiesto di scegliere tra due immagini e in base alla risposta data il campione è stato diviso in due gruppi casuali. Dopo è stata assegnata ai due rispettivi gruppi una domanda con due alternative. La prima alternativa era: "Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%?" (V15), mentre la seconda era: "Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 40%? (V16_2). La risposta a questa domanda era di tipo dicotomico, sì o no. Successivamente veniva fatta a tutti una seconda domanda a risposta aperta in cui veniva chiesto di indicare, in cifre, la percentuale giusta (V17).

Questa sequenza di domande è stata fatta per vedere quanto le persone sarebbero state condizionate nel rispondere alla seconda domanda a seconda del fatto che nella prima domanda ci fosse come valore il 10% o il 40%. Attraverso questa seconda domanda, infatti, si può individuare se c'è stato o meno l'uso dell'euristica dell'ancoraggio. Questo verrà fatto vedendo quanto sono diverse le risposte alla domanda aperta, di coloro che si sono ancorati al 10% rispetto coloro che si sono ancorati al 40%.

Statistics				
		Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%?	Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 40%?	Percentuale ONU
N	Valid	322	154	517
	Missing	229	397	34

Tabella 19 Numero di persone che hanno risposto alla domanda V15 o alla domanda V16_2. Fonte: dataset "domandevarie 2018-22".

Possiamo notare (Tabella 19) come la maggior parte del campione ha scelto la prima immagine e quindi gli è stata assegnata come percentuale di confronto, nella domanda successiva, il 10%. Infatti, alla prima domanda le risposte sono state 322, alla seconda 154.

Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	si	176	31.9	54.7	54.7
	no	146	26.5	45.3	100.0
	Total	322	58.4	100.0	
Missing	System	229	41.6		
Total		551	100.0		

Tabella 20 Frequenze delle risposte alla domanda V15. Fonte: dataset "domandevarie 2018-22".

Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 40%?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	si	51	9.3	33.1	33.1
	no	103	18.7	66.9	100.0
	Total	154	27.9	100.0	
Missing	System	397	72.1		
Total		551	100.0		

Tabella 21 Frequenze delle risposte alla domanda V16_2. Fonte: dataset "domandevarie 2018-22".

Prima di tutto per rispondere al quesito b e vedere se è stata usata l'euristica dell'ancoraggio bisogna andare a vedere quali persone sono state coerenti o no. Le persone coerenti sono quelle che alla domanda: "Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%?" hanno risposto "sì è superiore al 10%" o "no non è superiore di 10%" e quando veniva chiesto di indovinare il valore hanno scritto rispettivamente o un valore maggiore del 10% o minore del 10%.

Percentuale ONU * Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%? Crosstabulation				
		Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%?		Total
		si	no	
Percentuale ONU	Count			
		si	no	
	11	1	0	1
	12	2	0	2
	13	3	0	3
	14	2	0	2
	14	2	0	2
	15	15	0	15
	16	1	0	1
	16	1	0	1
	17	3	0	3
	18	2	0	2
	20	38	2	40
	22	1	0	1
	23	3	0	3
	25	9	0	9
	28	2	0	2
	30	19	2	21
	32	1	0	1
	33	1	0	1
	34	0	1	1
	35	6	0	6
	40	8	0	8
	50	7	0	7
	57	1	0	1
	60	13	0	13
	67	2	0	2
	70	4	0	4
	80	4	0	4
	85	1	0	1
	90	3	0	3
	95	1	0	1
	97	1	0	1
	98	1	0	1
	100	2	0	2
Total		160	5	165

Tabella 22 Percentuale scritta da chi aveva risposto alla domanda V15. Fonte: dataset "domandevarie 2018-22".

Nella tabella 22 possiamo vedere che nel caso della domanda “Secondo te la quota dei paesi africani partecipanti all’ONU è maggiore del 10%” cinque individui hanno risposto “no non è superiore del 10%” e successivamente hanno indicato una percentuale ONU superiore al 10%. Quindi ci sono state 5 persone che sono state incoerenti.

Invece nel caso della domanda “secondo te la quota dei paesi africani partecipanti all’ONU è maggiore del 40%” gli incoerenti sono stati 2 (Tabella 23). Infatti, due individui che avevano risposto di no poi hanno indicato delle percentuali più alte di 40%.

Percentuale ONU * Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 40%? Crosstabulation				
Count		Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 40%?		Total
		si	no	
Percentuale ONU	42	1	0	1
	45	6	0	6
	50	5	1	6
	55	1	0	1
	56	1	0	1
	57	1	0	1
	60	10	0	10
	65	1	0	1
	70	3	1	4
	72	1	0	1
	75	1	0	1
	80	2	0	2
	85	1	0	1
	90	5	0	5
	99	1	0	1
	100	2	0	2
Total		42	2	44

Tabella 23 Percentuale (V17) scritta da chi aveva risposto alla domanda V16_2. Fonte: dataset “domandevarie 2018-22”.

Statistics		
coerenti		
N	Valid	429
	Missing	122
Mode		1.00
Range		.00
Minimum		1.00
Maximum		1.00

coerenti					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	429	77.9	100.0	100.0
Missing	System	122	22.1		
Total		551	100.0		

Tabella 24 Numero di rispondenti che sono stati coerenti nel rispondere alle domande. Fonte: dataset "domandeverie 2018-22".

Attraverso ulteriori analisi (Tabella 24), con la creazione di una variabile "coerenti" siamo riusciti ad ottenere il numero dei rispondenti coerenti, che sono stati in tutto 429. Infine, tenendo conto solamente dei coerenti, si può guardare se è stata o meno usata l'euristica dell'ancoraggio.

Percentuale ONU * Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%?				
Percentuale ONU				
Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%?				
	Mean	N	Std. Deviation	
si	34.45	161	22.536	
no	5.69	128	2.022	
Total	21.71	289	22.106	

Tabella 25 Media delle percentuali (V17) date dai coerenti della domanda V15. Fonte: dataset "domandeverie 2018-22".

Percentuale ONU * Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 40%?				
Percentuale ONU				
Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 40%?				
	Mean	N	Std. Deviation	
si	64.23	44	17.770	
no	21.80	96	8.805	
Total	35.14	140	23.267	

Tabella 26 Media delle percentuali (V17) date dai coerenti della domanda V16_2. Fonte: dataset "domandeverie 2018-22".

Dalle analisi emerge (Tabella 25) a chi è stata fatta la domanda: "Secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 10%?" tra chi ha risposto sì o no a questa domanda e poi, in modo coerente, ha rispettivamente indicato un valore maggiore o minore

del 10%, ha risposto: il 21% dei paesi.

Invece (Tabella 26) coloro a cui è stata fatta la domanda: "secondo te la quota dei paesi africani partecipanti all'ONU è maggiore del 40%?" e hanno risposto di sì o di no, poi coerentemente con la risposta data, hanno indicato come valore un valore rispettivamente maggiore o minore del 40% che in media: il 35% dei paesi.

Quindi possiamo vedere che è plausibile che il campione abbia risposto alla domanda 16 facendo ricorso all'euristica dell'ancoraggio. In quanto a seconda del numero percentuale indicato nella prima domanda i partecipanti hanno risposto alla domanda aperta che chiedeva di scrivere la percentuale di paesi africani presenti all'ONU. A coloro cui è stato chiesto "Secondo te la quota di paesi africani presenti all'ONU è maggiore del 10%?" hanno risposto, al momento di scrivere una percentuale, delle risposte la cui media è 21,71%. Mentre a chi è stato chiesto "Secondo te la quota di paesi africani presenti all'ONU è maggiore del 40%?" al momento di indovinare la percentuale hanno dato risposte la cui media è stata del 35,14%. Questo fenomeno accade perché le percentuali date inizialmente funzionano come ancore che portano l'intervistato ad avvicinarsi a quei valori al momento di scegliere una percentuale. L'essere umano infatti tende ad ancorarsi a un fatto o ipotesi già conosciuta, anche se appena letta, e utilizzarla come punto di riferimento per scelte future, come può essere rispondere alla domanda seguenti.

ESERCIZIO 4

Punto a

Utilizzando i dati BES [3] sono stati scelti i seguenti domini:

1. Salute: indice di salute mentale (femminile)
2. Istruzione/formazione: bambini di 0-2 anni iscritti al nido
3. Lavoro e conciliazione dei tempi di vita: rapporto tra i tassi di occupazione (25-49 anni) delle donne con figli in età prescolare e delle donne senza figli
4. Benessere soggettivo: soddisfazione per la propria vita (femmine)
5. Benessere economico: benessere economico familiare

L'obiettivo dell'analisi è creare un indice composito in grado di misurare il benessere generale concentrandosi su quello femminile e considerando la sfera lavorativa, familiare ed economica.

Per creare l'indice composito la prima cosa che è stata fatta è stata individuare cinque domini di interesse e successivamente da questi sono stati selezionati degli indici.

Dal dominio della salute è stato scelto l'indice di salute mentale femminile, ritenuto un aspetto direttamente collegato al benessere femminile. Dal dominio dell'istruzione e della formazione è stato analizzato il tasso di bambini di 0-2 anni iscritti al nido, perché da quest'ultimo indice si può dedurre la capacità di una famiglia, indipendentemente dalla struttura del nucleo familiare, di poter accedere ai servizi di istruzione prescolare. Questa possibilità è data dalla disponibilità economica familiare e da un sistema di welfare che permette alle famiglie di usufruire di tali servizi. Infatti, dall'analisi ci si aspetta di ottenere risultati che indichino che più alta è la percentuale di bambini iscritti al nido più è alta anche la conciliazione della vita lavorativa e familiare. Inoltre, se la nascita di un figlio coincide con l'assenza di asili nido accessibili sono spesso le donne che devono rinunciare al lavoro per occuparsi esclusivamente al lavoro di cura del figlio. Perciò analizzeremo il rapporto tra i tassi di occupazione delle donne con figli in età prescolare e delle donne senza figli. Infine, verranno analizzate dal dominio del benessere soggettivo e da quello economico rispettivamente la soddisfazione per la propria vita e il benessere economico familiare, quest'ultimo è collegato strettamente al secondo e terzo dominio.

Il metodo di sintesi utilizzato è stato "approccio cardinale: campi di variazione", che è caratterizzato dal suo essere svincolato dall'unità di misura. Questo lo rende preferibile per domini, come in questo specifico caso, dove le unità di misura degli indici sono discordi. Dunque, escludendo "approccio cardinale: somma dei valori sommati in percentuali", rimaneva "approccio cardinale: scarti standardizzati", "approccio cardinale: campi di variazione" e "approccio cardinale: base fissa". Per semplicità di calcolo ed efficacia è stato scelto il secondo approccio tra quelli ammissibili.

Poiché riteniamo che i cinque domini abbiano tutti la stessa importanza non peseremo in maniera differente nessun dominio ma peseranno tutti uguali.

2021		
Regione	Sintesi	Posizione
Piemonte	0,4678	8
Valle d'Aosta/Vallée d'Aoste	0,5888	18
Liguria	0,4867	11
Lombardia	0,4934	12
Trentino-Alto Adige/Südtirol	0,7005	21
Provincia Autonoma di Bolzano/Bozen	0,6857	20
Provincia Autonoma di Trento	0,7268	22
Veneto	0,5833	17
Friuli-Venezia Giulia	0,5027	13
Emilia-Romagna	0,5626	16
Toscana	0,6191	19
Umbria	0,3522	6
Marche	0,4732	10
Lazio	0,4706	9
Abruzzo	0,4139	7
Molise	0,5458	15
Campania	0,1746	1
Puglia	0,2530	2
Basilicata	0,2621	3
Calabria	0,3185	5
Sicilia	0,3126	4
Sardegna	0,5306	14

*Tabella 27 Classifica dei risultati dell'indicatore composito (con 5 domini) per le regioni italiane in studio, nel 2021.
Fonte: dati BES.*

Dagli esiti dell'analisi (Tabella 27) risulta che alla Provincia Autonoma di Trento è associato il valore più alto dell'indicatore composito, cioè 0.7268. Ad essa seguono il Trentino - Alto Adige con un valore di 0.7005 e la Provincia Autonoma di Bolzano con un valore di 0.6857.

Tra le province con un indice più basso si osservano Campania con un valore di 0.1746, Puglia con un valore del 0.2530 e Basilicata con 0.2621.

I risultati dell'indicatore sono stati coerenti con quello che ci si aspettava, infatti nelle province autonome di Trento e Bolzano, dove è stata rafforzata l'offerta di servizi per l'infanzia, che sono di aiuto ai genitori per conciliare l'attività lavorativa con gli impegni familiari, il punteggio è più alto. [4]

Punto b

Si valutano adesso i cambiamenti tra il 2011 e il 2021 di 3 domini, scelti tra quelli considerati al punto precedente nelle regioni italiane, attraverso un indice composito formato dagli indicatori sintetici di ogni dominio. I tre domini analizzati sono: istruzione e formazione, lavoro e conciliazione dei tempi di vita, benessere economico. (Per il dominio del lavoro e della

conciliazione di tempi di vita, per assenza di dati del 2011, sono stati utilizzati i primi dati disponibili cioè quelli del 2016.

Anche in questo caso sia nel 2011 (Tabella 28) che nel 2021 (Tabella 29) la Provincia Autonoma di Trento è stata la regione con il punteggio più alto con un valore, rispettivamente, di 5.8651 nel 2011 e di 5.5687 nel 2021. In questa regione, pur mantenendo la prima posizione, con il passare del tempo il valore dell'indice composito è peggiorato leggermente.

REGIONI	SINTESI 2011	POSIZIONE	REGIONI	SINTESI 2021	POSIZIONE
Piemonte	4,6949	11	Piemonte	4,7882	15
Valle d'Aosta/Vallée d'Aoste	5,3836	19	Valle d'Aosta/Vallée d'Aoste	5,0623	20
Liguria	4,8373	12	Liguria	4,5833	8
Lombardia	4,9866	13	Lombardia	4,7040	11
Trentino-Alto Adige/Südtirol	4,9920	14	Trentino-Alto Adige/Südtirol	4,9264	16
Provincia Autonoma di Bolzano/Bozen	4,1433	8	Provincia Autonoma di Bolzano/Bozen	4,3436	7
Provincia Autonoma di Trento	5,8651	22	Provincia Autonoma di Trento	5,5687	22
Veneto	5,1887	17	Veneto	5,0446	19
Friuli-Venezia Giulia	5,3432	18	Friuli-Venezia Giulia	4,7675	14
Emilia-Romagna	5,4542	21	Emilia-Romagna	4,9509	17
Toscana	5,1801	16	Toscana	5,3055	21
Umbria	5,4064	20	Umbria	4,7280	12
Marche	4,5679	9	Marche	4,9902	18
Lazio	4,6436	10	Lazio	4,6052	9
Abruzzo	3,7435	6	Abruzzo	4,0900	6
Molise	4,1387	7	Molise	4,7546	13
Campania	2,5690	1	Campania	3,2294	1
Puglia	3,4399	3	Puglia	3,8644	4
Basilicata	3,5774	5	Basilicata	3,6696	3
Calabria	3,3442	2	Calabria	3,3950	2
Sicilia	3,4751	4	Sicilia	4,0003	5
Sardegna	5,0252	15	Sardegna	4,6283	10

Tabella 28 Classifica dei risultati dell'indicatore composito (con 3 domini) per le regioni Italiane in studio, nel 2011. Fonte: dati BES.

Tabella 29 Classifica dei risultati dell'indicatore composito (con 3 domini) per le regioni Italiane in studio, nel 2021. Fonte: dati BES.

Questo non accade per le regioni del Sud Italia, che sebbene mantengano l'ultima posizione con il passare degli anni non hanno peggiorato ulteriormente il valore dell'indice ma al

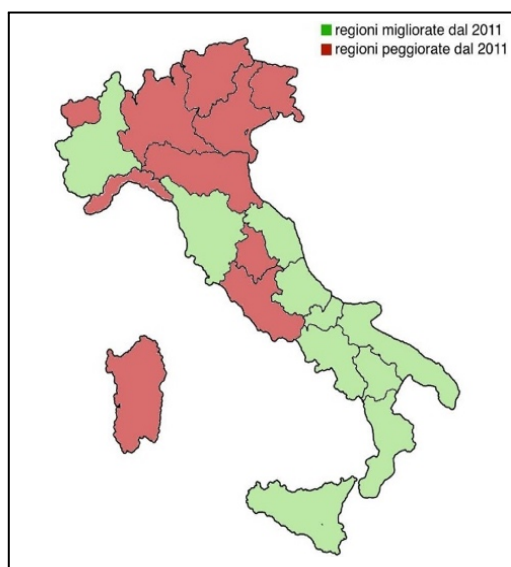


Grafico 9 variazione delle regioni italiane dal 2011 al 2021. Fonte: dati BES

contrario è migliorato. Tendenzialmente sono migliorate le regioni del Sud e sono peggiorate le regioni del Nord-Est (Grafico 9). Campania e Molise sono quelle che sono migliorate di più, mentre Umbria e Friuli-Venezia Giulia sono quelle che sono peggiorate di più. In Umbria la percentuale di bambini di 0-2 anni iscritti al nido è diminuita di quasi due punti percentuali, mentre in Friuli-Venezia Giulia è diminuita del 1,4%. Invece in Campania sia la percentuale di bambini di 0-2 anni iscritti al nido che il benessere economico è aumentato di quasi un punto percentuale. Le regioni che hanno subito minore variazione sono Lazio e Calabria.

L'indice di sintesi che è variato maggiormente è la percentuale di bambini di 0-2 anni iscritti al nido, con una variazione che passa da meno 2 punti percentuali a più 2.5 punti percentuali. Mentre l'indice che è variato di meno è il rapporto tra i tassi di occupazione tra donne con e senza figli, che è cambiato di poco più di mezzo punto percentuale sia in negativo che in positivo.

Punto c

Si valuta la concordanza al 2021 tra i 3 indicatori sintetici individuati al punto b e l'indicatore composito del punto a.

RHO	S1	S2	S3	S4
S1	1	0,59175607	0,470920384	0,7368718238
S2		1	0,6883116883	0,3083003953
S3			1	0,1027667984
S4				1

Rho Medio	0,4831545266
W	0,5692954389

Tabella 30 Concordanza tra i 3 indicatori sintetici individuati al punto b e l'indicatore composito del punto a, nel 2021. Fonti: dati BES

Dalla tabella 30 si nota la presenza della concordanza, nonostante non sia particolarmente elevata (0.5693). Essa viene abbassata dalla bassa correlazione tra il benessere economico e il lavoro e conciliazione dei tempi di vita (0.1027) e dalla bassa correlazione tra benessere economico e la percentuale di bambini iscritti al nido (0.3083). Osserviamo invece un'alta correlazione tra benessere economico e l'indicatore composito e tra il lavoro e conciliazione dei tempi di vita e la percentuale di bambini iscritti al nido (rispettivamente 0.7368 e 0.6883).

ESERCIZIO 5

Dagli obiettivi presenti nell'agenda 2030, nella seguente analisi, è stato studiato il goal 4: "Garantire un'istruzione di qualità inclusiva ed equa e promuovere opportunità di apprendimento permanente per tutti". [5]

Grazie ai dati forniti dalle Nazioni Unite [6] sono stati presi i valori di tre indicatori per ogni paese del mondo, in un range di anni che va dal 2017 fino all'anno più recente, il 2021, con i quali è stato creato un indicatore composito in grado di misurare l'approccio della popolazione all'istruzione in relazione ai fondi stanziati.

Data la presenza di valori mancanti per alcune aree geografiche in certi anni, sono stati considerati gli anni subito precedenti in modo da poter condurre comunque l'analisi. Invece i paesi che non presentavano dati in nessuno degli indici scelti sono stati elisi dallo studio.

Gli indicatori in studio:

4.1 Entro il 2030, garantire che tutte le ragazze e i ragazzi completino gratuitamente, istruzione primaria e secondaria equa e di qualità che portino a risultati di apprendimento pertinenti ed efficaci.

- **4.1.2** Tasso di completamento (istruzione primaria, inferiore, istruzione secondaria, istruzione secondaria superiore)

4.b Entro il 2020, espandere sostanzialmente a livello globale il numero di borse di studio disponibili per i paesi in via di sviluppo, in particolare paesi meno sviluppati, piccoli Stati insulari in via di sviluppo e Paesi africani, per l'iscrizione all'istruzione superiore, compresa la formazione e l'informazione professionale e tecnologia delle comunicazioni, tecnica, ingegneria e programmi scientifici, in sviluppo.

- **4.b.1** Volume dei flussi di aiuti pubblici allo sviluppo per le borse di studio per settore e per tipologia di studi

4.2 Entro il 2030, garantire che tutte le ragazze e i ragazzi abbiano accesso allo sviluppo della prima infanzia di qualità, cura e scuola materna istruzione in modo che siano pronti per l'istruzione primaria.

- **4.2.2** Tasso di partecipazione all'apprendimento organizzato (un anno prima dell'età ufficiale di ingresso primario)

Le informazioni dei tre indicatori scelti sono state filtrate in modo tale che per ogni paese, venga selezionata tutta la popolazione (quantile: _T), senza fare distinzioni di sesso (sex: bothsex) e con un livello di educazione limitato alla scuola primaria (Education Level: Primary). I dati risultano dunque filtrati in modo da poter avere un solo valore per ogni paese.

I paesi che presentano i dati necessari per l'analisi e che quindi sono stati considerati per il calcolo dell'indicatore composito sono 83. Assumendo che tutti gli indicatori abbiano la stessa rilevanza è stato scelto di non assegnare pesi differenti. La scelta del metodo di sintesi per la costituzione dell'indicatore composito è stata quella del "campo di variazione".

Per osservare visivamente le aree geografiche d'interesse, a seguito è riportata la mappa dei paesi selezionati (Mappa 1).



Mappa 1: paesi che rientrano nell'analisi

L'esito del metodo di sintesi applicato all'indicatore composito fornisce i seguenti risultati.

POSIZIONE	PAESI	SINTESI	POSIZIONE	PAESI	SINTESI
1	Indonesia	0,9700	43	Nepal	0,5524
2	Republic of Moldova	0,9677	44	Suriname	0,5452
3	Viet Nam	0,8175	45	Cambodia	0,5419
4	India	0,7567	46	Laos	0,5382
5	Brazil	0,7400	47	Dominican Republic	0,5367
6	Türkiye	0,7373	48	Guatemala	0,5250
7	Mongolia	0,7261	49	Honduras	0,5215
8	South Africa	0,7036	50	Egypt	0,5075
9	Serbia	0,7018	51	Namibia	0,4685
10	Mexico	0,7007	52	Armenia	0,4582
11	Thailand	0,6995	53	Samoa	0,4544
12	Belarus	0,6884	54	Benin	0,4475
13	Colombia	0,6862	55	Zimbabwe	0,4471
14	Peru	0,6786	56	Timor-Leste	0,4331
15	Argentina	0,6682	57	Papua New Guinea	0,4320
16	Chile	0,6649	58	North Macedonia	0,4307
17	Tonga	0,6596	59	Tanzania	0,4101
18	Saint Lucia	0,6571	60	Cameroon	0,4054
19	Cuba	0,6564	61	Rwanda	0,4011
20	Costa Rica	0,6515	62	Congo	0,3782
21	Fiji	0,6469	63	Belize	0,3732
22	Kyrgyzstan	0,6450	64	Gambia	0,3620
23	Kazakhstan	0,6427	65	Sudan	0,3537
24	Kiribati	0,6426	66	Tajikistan	0,3490
25	Maldives	0,6418	67	Myanmar	0,3457
26	Bangladesh	0,6403	68	Lesotho	0,3324
27	Bolivia	0,6368	69	Comoros	0,3185
28	Algeria	0,6294	70	Ethiopia	0,3055
29	Tuvalu	0,6264	71	Sierra Leone	0,3003
30	Albania	0,6252	72	Madagascar	0,2979
31	Ecuador	0,6121	73	Guinea	0,2891
32	Ghana	0,6105	74	Djibouti	0,2650
33	Philippines	0,5998	75	Burundi	0,2566
34	Uruguay	0,5963	76	Mali	0,2458
35	Uzbekistan	0,5938	77	Liberia	0,2330
36	Montenegro	0,5937	78	Côte d'Ivoire	0,2303
37	Togo	0,5893	79	Democratic Republic of the Congo	0,2025
38	Panama	0,5818	80	Senegal	0,1774
39	State of Palestine	0,5649	81	Niger	0,0989
40	Paraguay	0,5614	82	Central African Republic	0,0659
41	Jordan	0,5596	83	Chad	0,0447
42	Pakistan	0,5570			

Tabella 31 Classifica dei risultati dell'indicatore composito per i paesi che rientrano nell'analisi. Fonte: dati United Nations - Statistics - SDG Indicators Database

Al primo posto troviamo l'Indonesia, e a seguire Moldova e Vietnam mentre alle ultime posizioni troviamo le nazioni del Chad, Repubblica Centrafricana e Niger (Tabella 31).

Per riuscire a dare un'interpretazione all'indicatore composito abbiamo inizialmente analizzato in modo isolato i valori dei tre indici.

L'indice 4.1.2 relativo al tasso di completamento della scuola primaria si aggira su percentuali molto alte per quasi tutti i paesi sotto esame. Ad esclusione della Liberia, dove con la percentuale più bassa di completamento della scuola primaria rispetto a tutti gli altri paesi, è pari al 29%. A seguire hanno valori ancora molto bassi il Chad, la Repubblica Centrafricana, Niger e Mali dove solo dal 30% al 40% di popolazione completa la scuola primaria. Mentre per tutti gli altri paesi sotto analisi almeno il 50% della popolazione completa la scuola primaria.

L'indice 4.b.1 riguarda la quantità di aiuti con borse di studio date ai paesi poveri o in via di sviluppo per studiare nei paesi benefattori. Si suppone che maggiore sia il valore dell'indice e quindi maggiori siano gli aiuti ricevuti, maggiore è la difficoltà del paese a garantire l'istruzione ai propri cittadini. Indonesia, Moldova e India sono i paesi che ricevono più fondi e aiuti pubblici, hanno ricevuto rispettivamente 58, 53 e 28 milioni di dollari.

I valori più bassi in assoluto di questo indice li hanno Belize, Lesotho e Saint Lucia con rispettivamente il 0.21, 0.33 e 0.56 milioni di dollari, si suppone che i pochi aiuti siano anche proporzionali alle dimensioni territoriali e alle dimensioni della popolazione. Per confermare o confutare questa ipotesi bisognerebbe analizzare più dettagliatamente il processo e la scelta dei paesi a cui vengono destinati tali aiuti, questa analisi tuttavia non verrà svolta in questo studio.

Per l'indice 4.2.2, cioè il tasso di partecipazione all'apprendimento organizzato, i valori più alti sono relativi all'Argentina con un 100% e il Vietnam con un 99.91%, mentre i valori più bassi sono attribuiti a Myanmar (11.77%) e Tajikistan (12.50%).

Quindi risulta molto significativo per la prima posizione dell'Indonesia il valore dell'indice 4.b.1 dove ha il valore più alto di tutti i paesi (58.03 milioni di dollari), ma anche i valori degli indici 4.1.2 e 4.2.2 sono comunque alti, rispettivamente il 97% completa la scuola primaria e il 96% partecipa all'apprendimento organizzato.

La Moldova deve la sua seconda posizione principalmente per il valore dell'indice 4.b.1, essendo l'unica nazione assieme all'Indonesia a ricevere più di 50 milioni di dollari.

Mentre il Chad e la Repubblica Centrafricana devono la loro ultima posizione al fatto che i valori degli indici del 4.1.2 e del 4.2.2 sono tra i più bassi.

Bibliografia

- [1] «e-learning UniMiB,» 2023. [Online]. Available: <https://elearning.unimib.it/course/view.php?id=43515>.
- [2] «e-learning UniMiB,» 2023. [Online]. Available: <https://elearning.unimib.it/course/view.php?id=43515>.
- [3] «Istat,» 2023. [Online]. Available: [https://www.istat.it/it/benessere-e-sostenibilit /la-misurazione-del-benessere-\(bes\)/gli-indicatori-del-bes](https://www.istat.it/it/benessere-e-sostenibilit /la-misurazione-del-benessere-(bes)/gli-indicatori-del-bes).
- [4] &. C. T. A. c. d. F. Billari, «Rapporto sulla popolazione. L'Italia e le sfide della demografia.,» Bologna, Il Mulino, 2021, p. 67.
- [5] «United Nations,» 2023. [Online]. Available: <https://sdgs.un.org/goals/goal4>.
- [6] «United Nations - Statistics - SDG Indicators Database,» 2023. [Online]. Available: <https://unstats.un.org/sdgs/dataportal/database>.